# Optically responsive switches

Pijper, Thomas

# Chapter 2

# An introduction into computational chemistry

*This chapter provides an introduction into the field of computational chemistry. It introduces basic concepts in quantum chemistry, explains the most popular theories for describing molecular systems, and discusses applications that can be useful to (physical) organic chemists. Computational chemistry will be used through this thesis to augment spectroscopic studies. It is also the main topic of chapter 7, in which it is used to study various molecular systems.*

## 2.1 Introduction

Computational chemistry can be a valuable tool to the (physical) organic chemist. It can be used to study the characteristics of individual molecules, interactions between molecules (such as bonding), and the response of molecules to external perturbations (such as an electric field). In this chapter, commonly used theories and methodologies in computational chemistry will be discussed. The purpose of this discussion is to introduce the concepts that are used in this thesis, but also to serve as a general introduction to computational chemistry that is geared towards organic chemists, hopefully sparking their interest in this field of chemistry and making them consider utilizing calculations in their own research. It is for this second purpose that, in order to provide a more complete view on computational chemistry, this chapter will discuss a few concepts that have not been used in the research discussed in this thesis.

Because this chapter is geared towards (physical) organic chemists, it has been attempted to make the explanation of each theory as intuitive as possible, minimizing the need for prior knowledge of quantum chemistry. In the process, the mathematical equations that underlie each theory will be avoided as much as possible. Instead, these can be found in various textbooks that have been written on this topic. Obviously, 'dumbing down' a theory has the drawback that people might misunderstand its scope, which can be dangerous as each theory has its limitations which one must understand prior to use. For this reason, this chapter will try explain how each theory can be applied and where its strong points and limitations lie. Readers who would like to learn more about the quantum chemical basis underlying each theory are encouraged to read *Introduction to Computational Chemistry* by Frank Jensen (2$^{nd}$ edition, 2007, Wiley & Sons).

It is unfortunately not possible to provide a complete overview of the field of computational chemistry – it is simply too big. For this reason, a few topics are not discussed. The most important of these are force field methods (which treat atoms by means of classical mechanics), semi-empirical methods (typically parameterized, simplified versions of Hartree–Fock theory), valence bond theory (which considers individual bonds instead of molecular orbitals), and solvation models (that model the effect of a solvent shell on the molecular system under investigation). Furthermore, as this chapter is focused towards (physical) organic chemists, it focusses primarily on studying individual molecules. For this reason, methods popular in solid state physics as well as methods for simulating ensembles of molecules (such as Monte Carlo methods and molecular dynamics)

will not be covered. Finally, relativistic effects will be largely ignored as they usually play only a minor role in organic chemistry.

The next section will introduce the basic approximations commonly made in computational chemistry and introduce one of the basic theories, namely Hartree–Fock theory.

## 2.2 Basic approximations and the Hartree–Fock method

### 2.2.1 Basic approximations

The key to computational quantum chemistry is to obtain the *wave function* of the system that is being investigated (Ψ). The wave function describes the quantum mechanical behavior of the system. When the wave function is assumed to be independent of time and when relativistic effects are ignored, it can be obtained by solving the Schrödinger equation:

$$\hat{H}\Psi = E\Psi$$

Obtaining the exact wave function of a system is however incredibly complex. Already, we have made two important approximations: we have assumed the wave function to be independent of time, and we have ignored relativistic effects. Still, the complexity of molecular systems requires us to introduce additional approximations if we want to be able to solve the Schrödinger equation. A third important approximation we will make is the *Born–Oppenheimer approximation*.[1] In this approximation, it is assumed that because nuclei move with much lower velocities than electrons, they can be considered to be stationary with respect to the electrons. This is a sensible assumption and it simplifies the problem even further – by fixing the positions of the nuclei, we do not have to account for the effect of the movement of nuclei on electrons. Furthermore, the Born–Oppenheimer approximation allows us to define a *potential energy surface* (PES) on which each point corresponds to a certain set of positions for the nuclei and where each such point has a unique electronic wavefunction. As an example, Figure 2.1 shows the PES for the dissociation of a diatomic molecule where the *x*-axis displays the positions of the nuclei (the bond) and the *y*-axis displays the energy of the system.

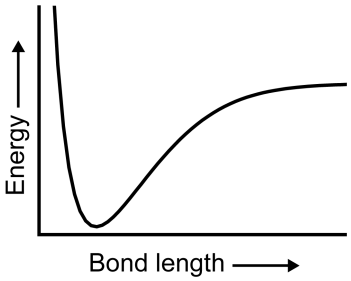


***Figure 2.1*** *An example of a one-dimensional potential energy surface for a diatomic molecule.*

A given system will have several of these surfaces, each corresponding to a unique electronic state. It is hereby also assumed that the wave function of our system is restricted to one PES. In other words, we assume that there is no coupling between surfaces (*i.e.* distinct electronic states). For this reason, the Born–Oppenheimer approximation is sometimes also referred to as the *adiabatic approximation*.[2]

We have now arrived at the point where we can solve the Schrödinger equation for a one-electron system such as $H_2^+$. Also, we are only one step away for solving it for a system containing more electrons. For an *n*-electron system, we can now construct the wave function from so-called *Slater determinants*,[3] mathematical expressions that describe the wave function in terms of molecular orbitals and the positions of electrons in these orbitals. The spin of each electron is also described, whereby the Pauli exclusion principle, which states that no two electrons may occupy the same quantum state, is obeyed. It should be noted that an exact description of Slater determinants is beyond the scope of this text, but they play an important role in the theories that will be discussed in the next few sections.

If we are now to construct the wave function starting from only one Slater determinant, we may be able to derive the *Hartree–Fock* method.

### 2.2.2 Hartree–Fock

Hartree–Fock is a theory that, despite some significant drawbacks, is very important in quantum chemistry. As said, the Hartree–Fock wave function is constructed from only one Slater determinant. This is actually yet another approximation, one that actually does not always give good results as will be discussed later on. The molecular orbitals are typically expressed in terms of a set

of known functions, the so-called *basis set* (one well-known example being the 6-31G set from Pople and co-workers). Basis sets will be explained later on in this chapter.

A wave function that satisfies the Hartree–Fock equations unfortunately cannot be constructed in one step as each individual orbital can only be determined when all other orbitals are known. Because of this complication, the Hartree–Fock procedure start with first 'guessing' a trial wave function. Then, starting from the trial wave function, the molecular orbitals are 'optimized' iteratively in order to improve the description of the wave function (Figure 2.2). Hereby, use is made of the so-called *variational principle*, which implies that the 'best' orbitals are the ones which give the lowest energy. Using this principle, the optimization procedure will search for the best set of orbitals by trying to lower the energy of the system as far as possible. For this reason, one will typically see the energy of the system decrease with each iteration.

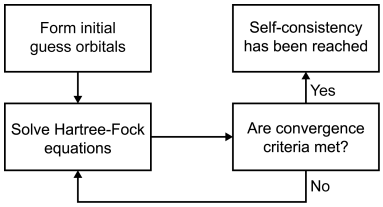


***Figure 2.2*** *Simplified representation of the SCF procedure.*

When the energy has reached a minimum, the orbitals (and therefore the wave function) are said to have reached *self-consistency* and the wave function calculation is said to have *converged*. The procedure itself is often referred to as *self-consistent field* (SCF). It should be noted that the term SCF is sometimes used as a synonym for Hartree–Fock, but this is not recommended as there are other methods that employ an iterative orbital optimization procedure.

### 2.2.3 Wavefunction types

With Hartree–Fock, different types of wave functions can be used, an overview of which is given in Figure 2.3. When the system is *closed shell*, having an even number of electrons which are all paired (in other words, a singlet system with only doubly occupied orbitals), a *restricted* wave function is normally used. In this type of wave function, each electron pair is forced to be in one orbital. If the system is *open shell*, it is possible to choose an *unrestricted* wavefunction.[4,5] Hereby, electrons of different spins (α and β) are allowed to occupy spatially different molecular orbitals. This gives the wave function more 'flexibility', *i.e.* it

might be able to describe systems for which a restricted wave function does not suffice. However, an unrestricted wave function has the drawback that it can be 'contaminated' by contributions from higher lying states. For example, a singlet state can be contaminated by triplet, quintet, etc. states which means that the wave function no longer corresponds to that of a pure singlet state. Similarly, a doublet state can be contaminated by quadruplet, sextet, etc. states. The amount of *spin contamination* in the wave function is given by the average value of $S^2$ for which the correct value can be calculated with the formula

$$\langle S^2 \rangle = S(S + 1)$$

where S is the angular spin momentum. Thus, for a doublet system, $\langle S^2 \rangle$ should equal 0.75. When, for example, a value of 0.76 is found, the amount of spin contamination is negligible. However, a significantly larger value indicates serious spin contamination. In such a case, an unrestricted wavefunction is not appropriate for describing the system and a different method should be used.

A second type of wave function that can be used for open shell systems is the *restricted-open shell* wave function,[6] in which all doubly occupied orbitals are described as with a restricted wave function. This wave function type is used less than the unrestricted wave function as it has the drawback that the converged solution has no unique set of orbitals. This has a few important consequences, such as that the orbital energies obtained by this method are virtually meaningless, and that methods that build upon the restricted-open Hartree–Fock wave function (such as those discussed in the next section) have to choose a unique set of reference orbitals.

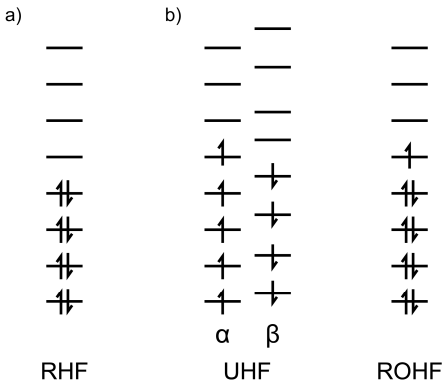


***Figure 2.3*** *a) A restricted wave function for a singlet system. b) An unrestricted and restricted-open wave function for a doublet system.*

### 2.2.4 Shortcomings of Hartree–Fock

So far, we have seen that by making a number of approximations, we are able to define the Hartree–Fock model. In addition, we have seen that the iterative optimization of the Hartree–Fock orbitals decreases the energy of the wave function till no further decrease is possible, at which point the calculation is considered to be converged. However, if we were to compare the energy obtained with Hartree–Fock with the energy corresponding to the exact solution to Schrödinger equation,[a] we would find that the Hartree–Fock energy is significantly higher. How is this possible? Are some of the approximations we made earlier perhaps incorrect?

Indeed, some of our approximations are incorrect. In principle, we are able to decrease the energy of our system further by using a basis set consisting of more functions, but even if we would keep increasing the size of our basis set, we would eventually reach a lower limit (Figure 2.4). This energetic limit is called the *Hartree–Fock limit*, and it is typically still quite far from the exact solution to the Schrödinger equation. Approaching the exact solution more closely requires that we examine the approximations made so far and decide to reject one or more of them.

The approximation that is the largest source of error in Hartree–Fock theory is the assumption that a single Slater determinant is able to give an accurate description of our system. This is generally not the case as the use of a single Slater determinant simplifies the way in which electrons interact with each other. Specifically, instead of considering each electron-electron interaction individually, Hartree–Fock only considers the interaction of each electron with a mean field of electrons. It is for this reason that the energy calculated with Hartree–Fock is too high, whereby the increase in energy due to this error in the Hartree–Fock model is generally referred to as the *electron correlation* energy. If this energy were to be added to the Hartree–Fock energy, the resulting energy would be identical to that of the exact solution to the Schrödinger equation (Figure 4).

[a] Actually, this 'exact solution' is not fully exact, as the time independency approximation and Born-Oppenheimer approximation still apply and relativistic effects are not considered. However, for the sake of simplicity we will from here on refer to this solution as the exact solution.

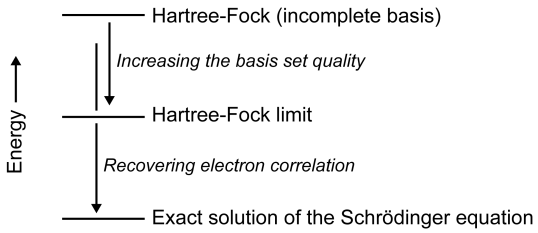


***Figure 2.4*** *The calculated energy can be brought closer to that of the exact solution by increasing the basis set quality. A method for recovering electron correlation energy is needed to get below the Hartree–Fock limit.*

The missing electron correlation is often seen as consisting of two parts. The *dynamic correlation* is the part missing due to not considering individual electron-electron repulsions. Several methods exist for recovering this part, though it is difficult to recover all of the missing dynamic correlation as this is computationally too expensive for all but the smallest of systems. Typically, only a part of the dynamic correlation is recovered, which already gives a marked improvement over results obtained with Hartree–Fock.

The *static correlation* is the part that is missing due to a fundamentally wrong description of the system by Hartree–Fock. This can happen in a number of cases, most prominently those where degeneracy plays a role. Consider heavy elements, which can have degenerate orbitals, or a case where two electronic states are energetically close to each other, such as in the vicinity of a conical intersection. An example of a small-sized system that displays degeneracy is the dissociation of the dihydrogen molecule. Figure 2.5 shows the dissociation PES as obtained at three levels: restricted Hartree–Fock (RHF), unrestricted Hartree–Fock (UHF), and the exact solution to the Schrödinger equation. As can be seen, RHF performs reasonable well around the equilibrium bond length. The energy at equilibrium might not be as low as that of the exact solution (due to missing dynamic correlation energy), but the shape of the PES is correct. However, at bond lengths over 1.25 Å, the shape of the PES begins to diverge from that of the exact solution. The reason for this is that, at this bond length, the two *1s* orbitals become degenerate. UHF may perform better than RHF because it can account for a limited amount of static correlation.

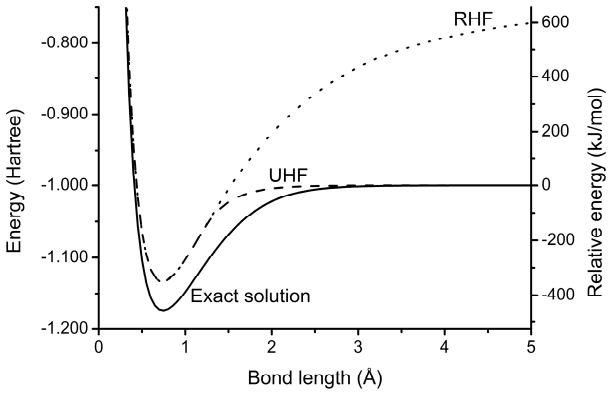


***Figure 2.5*** *$H_2$ dissociation curves, as obtained from restricted Hartree–Fock, unrestricted Hartree–Fock, and the exact solution to the Schrödinger equation.*

In the next two sections, various methods that account for electron correlation will be discussed. As these methods improve upon the Hartree–Fock model, they are commonly referred to as *post-Hartree–Fock* methods.

## 2.3 Recovering dynamic electron correlation

### 2.3.1 General considerations

As we have seen in our discussion of the Hartree–Fock model, the energy of a system as calculated with Hartree–Fock is generally too high because individual electron-electron repulsions are not considered. If we are to bring our energy closer to that of the exact solution to the Schrödinger equation, we need to recover this dynamic correlation energy.[b] The three most prevalent methods for achieving this are *configuration interaction*, *Møller–Plesset perturbation theory*, and *coupled cluster*. These methods share the characteristic that they generally start from a converged Hartree–Fock wave function. That makes sense as the Hartree–Fock description of the system is often already a good description – in many cases, it is only the dynamic correlation that is not accounted for. The three methods can therefore be seen as improving upon the Hartree–Fock model

[b] A critical reader may notice that our discussion on solving the Schrödinger equation is currently primarily focused on lowering the energy of our system. The reason for this lies with the variational principle, which states that a more accurate wave function always corresponds to a lower energy. Because of this principle, we can use a lowering of the energy as an indication that the description of the system is being improved (though only for theories that are variational).

instead of representing fundamentally different approaches to solving the Schrödinger equation.

As the short-comings of Hartree–Fock theory originate from its use of a single Slater determinant, the most obvious way to improve upon Hartree–Fock is to add more of these Slater determinants. The three methods for recovering electron correlation discussed herein work this way. Building on the Hartree–Fock reference Slater determinant, they construct additional determinants in which, with respect to the reference determinant, one or more electrons have been excited from an occupied orbital to a virtual orbital.[c] These *excited Slater determinants* are denoted by the number of electrons that has been excited. In *Singles* (S), only one electron has been excited; in *Doubles* (D), two electrons have been excited; in *Triples* (T), three electrons have been excited, and so on. Formally, there are no restrictions on the spin state – excited Slater determinants may be singlet, triplet, etc. However, only determinants with spin states corresponding to the spin state of the system will typically be used. Figure 2.6 provides a few examples of excited Slater determinants. Note that 'excitation' in this context only applies to additional Slater determinants that are used. It should not be confused with the electronic excitation of an atom or molecule.

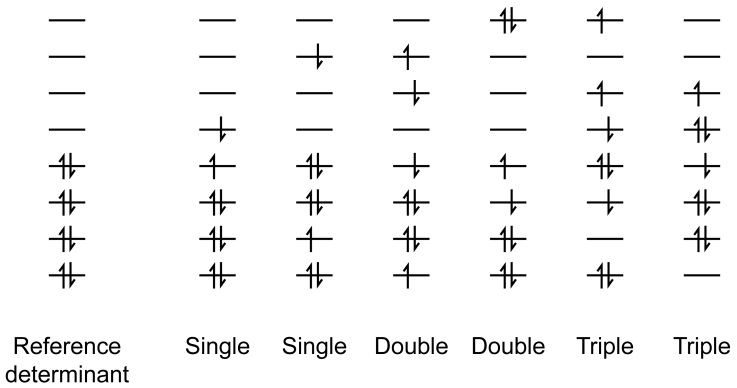


Figure 2.6 *Example of excited Slater determinants that can be constructed based on a reference determinant.*

What may be obvious at this point is that the excitation of electrons to virtual orbitals allows one to construct an incredibly large number of excited Slater

[c] In computational chemistry, unoccupied orbitals are called virtual orbitals. As molecular orbitals are constructed from basis functions, the number of virtual orbitals increases with the number of functions in the basis set (*i.e.* the size of the basis set).

determinants. The number of Singles (S) is still small, but the number of higher excited determinants increases factorially with each increase in excitation level. If all possible determinants were to be used in the post-Hartree–Fock calculation, both dynamic *and* static electron correlation would be completely taken into account (provided the basis set has an infinite amount of functions) and the wave function obtained would correspond to that of the exact solution (see also Figure 2.7). However, from a computational standpoint such a calculation would be too demanding for all but the smallest systems. This is especially true because post-Hartree–Fock methods require large basis sets to give accurate results, whereby a large number of basis functions leads to a large number of virtual orbitals (which in turn increases the number of excited determinants that can be constructed).

In order to limit the number of determinants, methods designed for recovering dynamic correlation typically limit the excitation level of the Slater determinants used in the calculation. In practice, the use of determinants with excitations higher than D already have a serious computational cost, though they can be used for small systems. Determinants with excitations above Q are rarely employed. In addition to limiting the excitation level, the number of determinants can be decreased further by not allowing excitations from certain orbitals. It is, for example, common not to excite electrons from the atomic cores (these core are then said to be *frozen* in the calculation), which is a sensible approximation since the core orbitals are typically unimportant in chemical reactivity.

As methods for recovering dynamic electron correlation often use a Hartree–Fock determinant as the reference, it is extremely important to note that these methods depend on the Hartree–Fock wave function already being a good description of the system. If this is not the case, for example when static electron correlation plays an important role (which is the case for systems involving degeneracy), these methods might not give sensible results. Nonsensical results might also be obtained when a spin-contaminated UHF wave function is used as a reference.

Finally, it should be noted that methods based on an ROHF reference determinant are possible and exist, but are more complex in their design as they have to choose a unique set of orbitals to use a reference (recall that ROHF does not provide a unique set of orbitals).

### 2.3.2 Configuration interaction

The configuration interaction (CI) method[7] is the oldest method for treating electron correlation and is, just as with Hartree–Fock, based on the variational principle. With CI, linear combinations of multiple Slater determinants are used to construct so-called *configurational state functions* (CSFs) which, in turn, are used to construct the wave function. As discussed earlier, a calculation that would include all possible excited Slater determinants (*i.e.* when there is no limit to the excitation level) would treat electron correlation fully and provide an answer identical to the exact solution. Such a calculation is called a *full CI* calculation and, as was also discussed, is typically too demanding as the amount of Slater determinants and, therefore CSFs, is too large. It is thus necessary to limit the excitation level. CI methods in which the excitation level is limited are commonly referred to as *truncated CI* methods. An overview of several common truncated CI methods and their relation to Hartree–Fock and the exact solution is given in Figure 2.7.

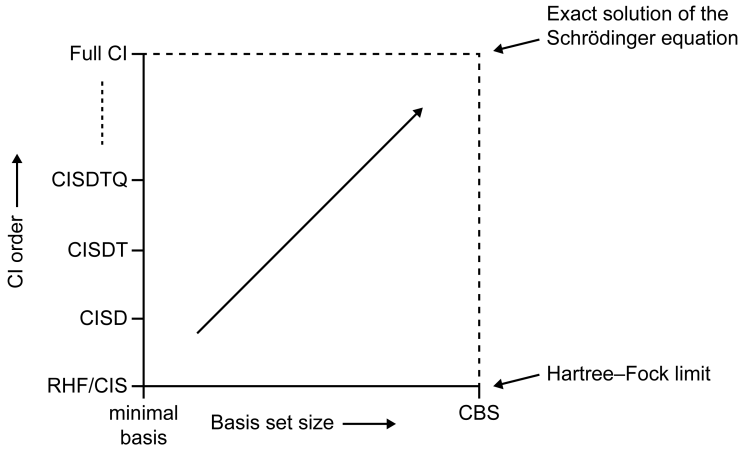


***Figure 2.7*** *The relation of the energy of various truncated CI methods with that of Hartree–Fock and the exact solution. 'CBS' stands for complete basis set, a set with an infinite amount of functions.*

When we limit the amount of determinants to single excitations, we obtain the *CI with Singles* (CIS) model. This model does not provide an improvement over Hartree–Fock with respect to the energy, but it does provide an improvement when calculating molecular properties (in which the change in wave function upon an external perturbation is assessed). In addition, CIS can be used for

'crude' excited state calculations.[d] Actually, we hereby touch upon an interesting characteristic of CI. The mathematical model behind CI makes it very easy to calculate excited states, which is not the case with the other two methods discussed in this section. However, it should be noted that since the reference orbitals are typically from a Hartree–Fock calculation of the ground state, the CI method can be considered to be 'biased' against excited states.

A marked improvement over Hartree–Fock is obtained when doubly excited determinants are used, which gives rise to the *CI with Doubles* (CID) method. We could hereby choose to include Singles as well (which in the presence of Doubles *do* provide an energetic improvement) as this has only a marginal impact on the total computational cost. This gives us the well-known CISD model. CISD typically accounts for a large amount of the dynamic correlation, and is able to include a small fraction of static correlation as well (due to the presence of Singles). If Triples are also included, the CISDT model is obtained, which recovers more electron correlation but has a much higher cost than CISD. Similarly, the inclusion of Quadruples gives the CISDTQ model, which gives results very close to those of a full CI calculation. In practice, CISD is the most commonly used method due to the high computational cost associated with CI. The computational cost of CISD generally scales as $N^6$ with an increase in the number of basis functions (where $N$ is the number of basis functions) while CISDT and CISDTQ scale as $N^8$ and $N^{10}$, respectively.

It should be mentioned that the CI method has two important drawbacks, which are best introduced by an example. Consider the case of a CID calculation on helium atoms. For a system consisting of a single helium atom, this means that determinants used are doubly excited with respect to the reference determinant. For a system consisting of two helium atoms, this would lead to determinants in which one helium atom is doubly excited, those in the other helium atom is doubly excited, and those in which both atoms are singly excited. However, determinants in which both atoms are doubly excited at the same time are not possible as such a configuration would count as a quadruple excitation. For this reason, the CID energy of two non-interacting helium atoms does not equal twice the CID energy of a single helium atom. This problem is commonly referred to as *size inconsistency*. More generally formulated, a method is size inconsistent when the calculated energy of a system of two (or more) non-interacting fragments is different from the energy that is obtained when these

[d] Here, we *are* talking about the electronic excitation of an atom or molecule.

fragments would be calculated separately. Truncated CI methods are all size consistent.

A second drawback to CI is *size inextensiveness*, which is a concept that is closely related to size inconsistency. It is said that a method is size inextensive when the energy of a system of two or more non-interacting fragments changes when the distance between these fragments is changed. It is a more difficult term to explain, but an important consequence of the size inextensiveness of CI is that CI recovers increasingly less electron correlation for increasingly large systems. It should be stressed that these two drawbacks only apply to truncated CI. Full CI is fully size consistent *and* size extensive.

### 2.3.3 Møller–Plesset perturbation theory

A second approach to recovering dynamic electron correlation is perturbation theory. In perturbation theory, it is assumed that a problem can be solved starting from a related problem that is already solved. Starting from a reference wave function, such a perturbation can be of the following form:

$$\hat{H} = \hat{H}_0 + \lambda\hat{V}$$

Here, $\hat{H}_0$ is the reference Hamiltonian, $\hat{V}$ is the perturbation to $\hat{H}_0$, $\lambda$ is the perturbation parameter which determines the strength of the perturbation, and $\hat{H}$ is the result of this perturbation. This perturbation scheme can be written as an expansion of the following form:

$$E = \lambda^0 E_0 + \lambda^1 E_1 + \lambda^2 E_2 + \lambda^3 E_3 + \ \ldots$$

where $E$ is the energy to be known, $E_0$ is the energy of the reference wave function, and $E_1$, $E_2$, and $E_3$ are the first-order, second-order, and third-order correction to the energy. In Møller–Plesset perturbation theory,[8] one of the most common applications of perturbation theory in quantum chemistry, it is attempted to obtain the energy of the exact solution to the Schrödinger equation by applying perturbations (usually) to a Hartree–Fock reference wave function.

The lowest Møller–Plesset correction order capable of recovering dynamic electron correlation is the second order, which is commonly referred to as the MP2. This method uses the first-order corrected wavefunction to recover an already large amount of dynamic correlation at a modest computational cost. Similarly to CI, MP theory makes use of excited Slater determinants; at the MP2

level, only contributions from Doubles are considered. MP2 scales approximately as $N^4$ to $N^5$ with an increase in the number of basis functions, which makes this method applicable to medium to large size systems.

The next MP order is MP3, which, as for MP2, also uses a first-order corrected wavefunction, but is able to recover more electron correlation. As for MP2 only Doubles are considered, however, the cost of MP3 is noticeably higher than that of MP2 as MP3 scales as $N^6$. In practice, MP3 usually does not provide a significant improvement over MP2, and as a consequence it is used less.

***Table 2.1*** *Overview of MP methods.*

| Møller–Plesset order | Determinants used | Uses *n*th-order Ψ |
|---|---|---|
| MP2 | D | first-order |
| MP3 | D | first-order |
| MP4 | S, D, T, Q | second-order |
| MP5 | S, D, T, Q | second-order |
| MP6 | S, D, T, Q, 5, 6 | third-order |
| MP7 | S, D, T, Q, 5, 6 | third-order |

A second-order correction to the wave function is used starting at the fourth MP order. This order, denoted as MP4, uses contributions from S, D, T, and Q determinants and scales as $N^7$. Alternatively, it is possible to exclude the T contributions (which are computationally the most demanding), which gives rise to the MP4(SDQ) model. MP4(SDQ) scales as $N^6$ and is moderately more expensive than MP3 while recovering significantly more correlation energy. The next step up, MP5, also uses the second-order corrected wave function, and uses the same determinants as MP4. This method, however, scales as $N^8$ and is therefore only suitable for small systems. Finally, MP orders beyond MP5 exist, but because of the high computational cost, MP5 and higher methods are rarely used. Table 2.1 provides an overview of MP levels.

From this discussion, it can be seen that an $n$th-order correction to the wave function allows for MP orders up to $2n+1$. In other words, after each two successive MP orders, a higher order corrected wave function is used for the next two MP orders. For some systems, this can result in an oscillatory behavior of the calculated energy when increasingly higher MP order methods are used. In Figure 2.8, this behavior is shown for the lithium fluoride molecule. It can be seen here that MP2 sort of 'overestimates' the amount of correlation energy. MP3 consequently reduces this amount, but with MP4 (which is the first MP order to use the second-order wave function) the amount of correlation energy is

again much higher, after which it is again somewhat decreased at the next MP order.

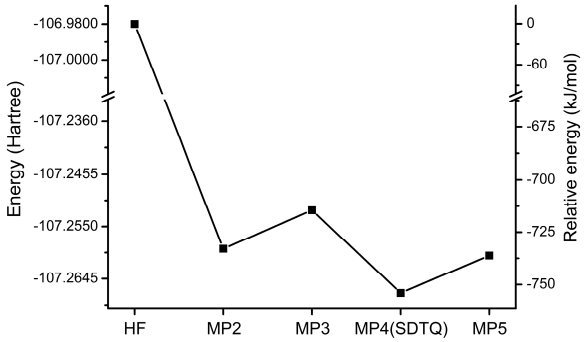


***Figure 2.8*** *Energies of a lithium fluoride molecule as calculated with the lowest order Møller–Plesset methods.*

The above behavior of Møller–Plesset theory highlights an important characteristic of this method: contrary to Hartree–Fock and configuration interaction, MP theory does not follow the variational principle. This means that increasingly higher MP orders may not give an increasingly lower energy. Actually, we hereby touch upon an important drawback of MP theory: there is no guarantee that the use of increasingly higher MP orders causes the convergence to a certain answer. For the lithium fluoride example in Figure 2.8 we do see the energy converge to a certain value, but this is not always the case. One of the reasons for this is that perturbation theory assumes the reference wave function is already a good description, and that the perturbation needed is thus only small in some sense. In cases where the Hartree–Fock wave function is far from the exact solution, for example in cases where static correlation plays in important role, MP theory can therefore lead to nonsensical results. This is to a large extent also true for configuration interaction, but while the use of more determinants in CI (CIS, CISD, CISDT, and ultimately full CI) will eventually cause one to reach the exact solution, this is not the case for MP theory. It is also for this reason that higher MP methods such as MP5 and MP6 find little use. CI methods (and CC methods, see below) with a similar cost often give more accurate results. It should be noted though that MP theory is size consistent as well as size extensive.

Finally, a commonly used empirical modification to MP theory should be mentioned. The electron correlation recovered in MP2 (and MP3) can in principle be categorized as belonging to one of two types: that originating from

opposite-spin interactions and that originating from same-spin interactions. Investigations into the deficiencies of MP2 have shown that the contribution from opposite-spin interactions is usually underestimated while that from same-spin interactions is usually overestimated. This has led to the development of the *spin component scaled* (SCS) MP2 method,[9] in which the individual spin contributions are scaled by empirically established parameters: 6/5 for the opposite spin correlation and 1/3 for the same spin correlation. SCS-MP2 often provides results more accurate than those of MP2, though the degree of improvement over MP2 fluctuates heavily with the type of system being calculated. Other parameters have been suggested,[10] leading to variations like SCSN-MP2[11] and SOS-MP2.[12] An SCS scheme has also been proposed for MP3, leading to the SCS-MP3 method[13] which, as of yet, has not been used extensively.

### 2.3.4 Coupled cluster

A third class of methods for treating dynamic correlation that will be discussed here are the coupled cluster (CC) methods.[14] CC methods can be seen as being related to Møller–Plesset theory – whereas MP theory uses excited Slater determinants of multiple types (S, D, T, etc.) to provide a correction of a given MP order, CC methods use only determinants of a given type to provide a correction of *all* orders. Like MP theory, CC methods do not follow the variational principle but are size consistent as well as size extensive. Generally speaking, CC methods outperform MP theory and truncated CI methods in terms of accuracy but do so against a higher computational cost.

As is the case for CI, CC methods do not improve upon Hartree–Fock when only contributions from Singles are included. Thus, the lowest level CC method that can be used is CCD. Inclusion of Singles consequently leads to the CCSD method, which scales as $N^6$ with the size of the basis set and which is computationally more demanding than CISD. The inclusion of Triples leads to the CCSDT method which, due to its very high computational cost, can be applied only to small systems. As with CI, contributions from increasingly higher excited determinants can be included in CC. If this were done, the resulting calculated wave function would approach that of the exact solution until at some point, when contributions from all determinants would be included, the exact solution would be reached.

Some variations on CC methods which attempt to lower the cost of CC calculations should be mentioned. One of the most important variations is to calculate the contribution of Triples and higher determinants perturbatively and

add these to the CC results. Various such methods exist, but the CCSD(T)[15] method is perhaps the most commonly used one. It should hereby be noted that a perturbative treatment of certain types of determinants is not unique to CC methods. This approach can (and has) also be applied to CI, leading to methods such as CIS(D).[16]

***Table 2.2*** *Overview of CC methods and related methods.*

| **D contributions** | **S,D contributions** | **S,D,T contributions** |
|---|---|---|
| CCD | CCSD | CCSDT |
| QCID | BD | BDT |
| | QCISD | QCISDT |
| **S contributions with perturbative D contributions** | **S,D contributions with perturbative T contributions** | **S,D,T contributions with perturbative Q contributions** |
| CCS(D) | CCSD(T) | CCSDT(Q) |
| | BD(T) | BDT(Q) |
| | QCISD(T) | QCISDT(Q) |

Methods that are very closely related to CC are *Brueckner theory* and *quadratic configuration interaction*. Brueckner theory optimizes the reference orbitals in such a way that contributions of Singles is exactly zero.[17,18] For this reason, *Brueckner Doubles* (BD) gives results similar to those of CCSD at a similar cost, while results of BD(T) are typically equivalent to those of CCSD(T). Quadratic configuration interaction (QCI) is a method that was developed to overcome the size inextensiveness of CI.[19] QCI methods are very similar to CC methods with respect to both results and cost. Typically, hardly any difference is found between results from CCSD(T), BD(T), and QCISD(T) calculations. Table 2.2 provides an overview of all CC methods and related methods discussed so far.

### 2.3.5 Concluding remarks

When comparing all post-Hartree–Fock methods discussed so far, they can roughly be ordered as follows, in order of increasing accuracy:

HF << MP2 < CISD < MP4(SDQ) ~ CCSD < MP4(SDTQ) < CCSD(T) < CCSDT

Here, each CC method can be substituted for a QCI or Brueckner method of the same level, as these typically give very similar results. As an increase in accuracy comes at the cost of increased computational requirements, the MP2 method is arguably the most interesting to an organic chemist. The more accurate methods

are typically too complex for an averaged sized molecular system, though a simple energy calculation might be feasible with some.

## 2.4 Multi-configurational SCF methods

### 2.4.1 General considerations

As mentioned in our discussion of the Hartree–Fock model, there are systems which cannot be described properly by a restricted Hartree–Fock wave function. Hereby, the error in the description is commonly referred to as static electron correlation. Most often, such difficult systems involve degenerate orbitals, the dihydrogen dissociation system of Figure 2.5 being an example of such a system. An unrestricted wave function might be able to provide a better description, but can only include a limited amount of static correlation. Furthermore, methods for recovering the dynamic correlation energy, for example truncated CI, often do not perform much better for these systems, though some of them *are* able to account for a limited amount of static correlation. However, the full CI method is able to account fully for both static and dynamic correlation. This raises a question: why is it that a truncated CI method such as CISD cannot fully account for static correlation while full CI can? Are there perhaps CSFs that are important for describing static correlation, but which are missing from truncated CI?

This is indeed the case, which brings us to the topic of *multi-configurational self-consistent field* (MCSCF) methods. MCSCF methods can be seen as a variation on truncated CI, where it is the determinants important for static correlation (instead of those important for dynamical correlation) that are included in the calculation. In addition, being SCF methods, these methods use an iterative optimization procedure that improves the orbitals. These methods are therefore able to handle very complex cases, but unfortunately also require a lot of skill to use. While the methods discussed so far can often be used as black box methods, this is not case for MCSCF methods as problems can arise with the selection of important CSFs, the SCF convergence procedure, and other aspects of MCSCF calculations.

### 2.4.2 CASSCF

As we now know that certain CSFs can be used to describing static correlation effects, the next question is: how do we select only those CSFs that are important? A very popular approach to this problem is the *complete active space SCF* method, commonly abbreviated as CASSCF.[20] Alternatively, this theory is also known as *fully optimized reaction space MCSCF* (FORS-MCSCF)[21] which differs

only in its technical implementation.[22] In CASSCF, molecular orbitals are partitioned into two spaces: an *active space* and an *inactive space* (Figure 2.9a). The active space contains all occupied and virtual orbitals that are important when accounting for static correlation and these are treated in a full CI calculation. The inactive space contains all orbitals which are considered not important, and these are treated in an RHF-like manner. This 'combination' of full CI with a Hartree–Fock like description makes CASSCF a variational method. In addition, CASSCF is both size consistent and size extensive, provided that the active space is chosen in such a way that its composition remains constant throughout all calculations.

As might be evident, the active space cannot include a large number of orbitals, otherwise the calculation would become too costly (see Table 2.3). An active space consisting of 14 electrons and 14 orbitals – which is denoted as CASSCF(14,14) – already requires a significant computational cost. For large systems consisting of many orbitals and electrons, this means that the largest part of the dynamic correlation energy will remain unaccounted for. Methods for recovering this dynamic correlation energy will be discussed later on in this section.

***Table 2.3*** *Active spaces size* vs *the number of singlet CSFs.*

| **Active space size (electrons, orbitals)** | **Number of singlet CSFs** |
|---|---|
| (2,2) | 3 |
| (4,4) | 20 |
| (6,6) | 175 |
| (8,8) | 1 764 |
| (10,10) | 19 404 |
| (12,12) | 226 512 |
| (14,14) | 2 760 615 |
| (16,16) | 34 763 300 |

In addition to describing static correlation effects, CASSCF is often also used for excited state calculations. The reason for this lies in its ability to optimize the molecular orbitals. Recall our earlier discussion of the CIS method. During this discussion, we noted that CI methods are able to calculate excited states, but that these calculations can be considered as 'biased' because the Hartree–Fock reference orbitals correspond to the ground state. Contrary to Hartree–Fock, CASSCF is able to optimize its orbitals not only for the ground state, but also for excited states. In addition, it is able to optimize its orbitals in such a way that

they provide for a balanced description of multiple electronic states as once, a procedure known as *state-averaged* CASSCF (SA-CASSCF). State-averaged orbitals are necessary when studying interactions between states, such as conical intersections or avoided crossings, as such studies usually require an equally good representation of the states involved. Some QC programs can also average the orbitals over states of different multiplicity, which is useful for example when studying spin-orbit coupling.

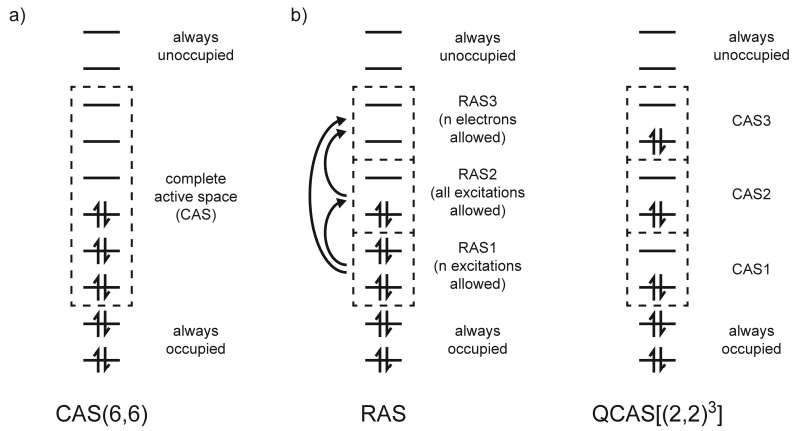


***Figure 2.9*** *a) Example of a complete active space consisting of 6 electrons in 6 orbitals. b) Active spaces employed in RASSCF and QCAS-SCF, two variations on CASSCF.*

### 2.4.3 Reducing the amount of CSFs

As noted, the size of the active space cannot be chosen to be too large, otherwise the calculation will become unmanageable. For this reason, several methods based on CASSCF have been developed that attempt to reduce the amount of CSFs. Two of such methods are depicted in Figure 2.9b. In *restricted active space SCF* (RASSCF),[23] the active space is partitioned into three restricted actives spaces: RAS1, RAS2, and RAS3. RAS1 has the restriction that only a limited number of electrons is allowed to be excited to RAS2 and RAS3. RAS2 has no restrictions, any number of its electrons may be excited to RAS3. RAS3 has the restriction that it may only contain a limited amount of electrons. In addition, any number of excitations is allowed within each RAS. This approach can be used to reduce the amount of CSFs considerably, but has the disadvantage that it is no longer size consistent and size extensive. A related method, named QCAS-SCF,[24] uses a different approach where the active space is reduced to a product of smaller active spaces where no excitations from one active space to any other active space are allowed.

One last approach worth mentioning is *occupation restricted multiple active space*, or ORMAS.[25] ORMAS allows one to set up any number of spaces, where each space has a certain minimum and maximum occupation number and excitations between spaces are allowed. This method thus provides much freedom in the choice of active space, and can be used to reproduce results of other methods such as RASSCF. Because of the flexibility of the method, it is not depicted in Figure 2.9. It should be noted that ORMAS can also be used as a CI only method, for which both Hartree–Fock and MCSCF wave functions can be used as the reference wave function.

### 2.4.4 Dynamic correlation in MCSCF

As we have seen earlier, dynamic correlation is important for an accurate description of the wave function. For this reason, various methods exist which attempt to recover this dynamic correlation, starting from a MCSCF reference wave function (Table 2.4). The simplest of these is *multi-reference CI* (MRCI), which typically starts at the MRCID or MRCISD level as considering Singles only does not recover much dynamic correlation. Needless to say, MRCI is a very costly method – the number of CSFs involved is roughly equal to the number of CSFs of the (non-MR) CI method multiplied by the number of CSF of the MCSCF reference wave function. This makes MRCI only applicable to very small systems. However, when it can be applied it yields a very accurate wave function. As with other truncated CI methods, MRCI is not size consistent and size extensive.

Because of the cost of MRCI, a more viable method for recovering the dynamic correlation is perturbation theory. The development of such a method is not trivial as a reference wave function has to be chosen prior to the perturbation (similar to that done for a ROHF wave function). Currently, several such perturbation theory methods are available, which differ mainly in their choice of reference wave function. Some of the more generally used methods are CASPT2,[26] MRMP2,[27] and NEVPT2,[28] which all apply a MP2-like correction to the reference wave function.

***Table 2.4*** *Popular methods capable of recovering dynamic correlation based on an MCSCF wave function.*

| Single-state PT2 | Multi-state PT2 | Configuration interaction |
|---|---|---|
| CASPT2 | MS-CASPT2 | MRCIS |
| MRMP2 | MCQDPT2 | MRCISD |
| NEVPT2 | QD-NEVPT2 | MRCISDT |
| | XMCQDPT2 | |
| | XMS-CASPT2 | |

Dynamic correlation is also important for the study of excited states. Not only the energy, but also the shape of the ground/excited state surfaces can sometimes not be described accurately by MCSCF alone. To this extent, multi-state MR-PT methods have been developed. These methods are able to use several CASSCF states in the formation of each 'real' MR-PT2 state, which makes them capable of describing highly challenging systems that are not well described by CASSCF alone. Well-known, related methods are multi-state CASPT2 (MS-CASPT2),[29] second-order multi-configurational quasi-degenerate PT (MCQDPT2),[30] and quasi-degenerate NEVPT2 (QD-NEVPT2),[31] which are the multi-state versions of the three (single-state) MRPT2 methods above. It should be noted that an improvement on the above multi-state methods was recently presented as the XMCQDPT2 method[32] (and also applied to MS-CASPT2 as XMS-CASPT2[33]).

An example of a system for which CASSCF does not work well is shown in Figure 2.10, which displays the dissociation curve of the lithium fluoride molecule. LiF is a very difficult case for QC methods (it is sometimes regarded as a 'benchmark' system). In the equilibrium geometry the bond in LiF is predominantly ionic, whereas the molecule dissociates into neutral atoms (which is typical for diatomic molecules in the gas phase). This means that, during the dissociation process, the first excited state (in which the bond is covalent) will decrease in energy and at some point 'cross' the surface that corresponds to the ionic bond. However, rather than actually crossing, the symmetry of the molecule causes the formation of an avoided crossing. As can be seen, SA-CASSCF predicts that this avoided crossing is located around 4.5 Å. The XMCQDPT2 method however places the avoided crossing around 6.5 Å and, in addition, predicts that the energy difference at this crossing is much smaller than was calculated by CASSCF. The surfaces calculated by XMQDPT2 are virtually identical to surfaces calculated by MRCISD.

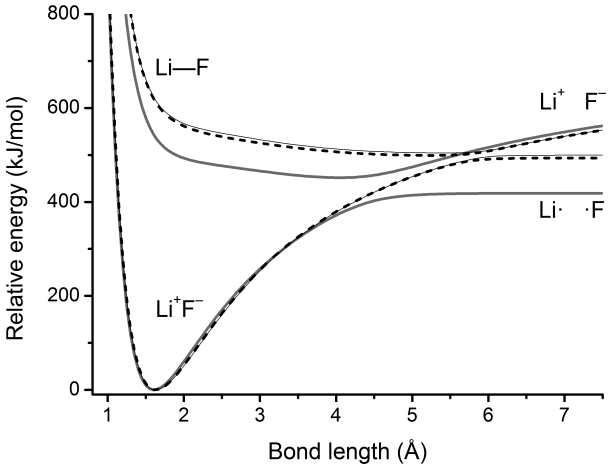


***Figure 2.10*** *PESs of the two lowest states for the dissociation of lithium fluoride, as calculated with SA-CASSCF (solid, grey), XMCQDPT2 (dashed), and MRCISD (solid, thin).*

## 2.5 Density Functional Theory

### 2.5.1 Describing a system in terms of electron density

A common factor of the methods discussed so far is that they provide information on the system being investigated by determining the wave function of the system. *Density functional theory* (DFT), another prevalent theory in computational chemistry, is based on a different approach. In DFT, it is assumed that the energy of a system is dependent completely on the *electron density* of that system (ρ). Hereby, 'functional' is a mathematical concept, namely a function that depends upon one or more variables which are functions themselves.

Several models have been suggested for how to use the electron density in calculating the energy of a system. The most important of these is *Kohn–Sham* (KS) theory,[34] a model that is closely related to Hartree–Fock theory. In KS, it is assumed that the kinetic energy of a system can be separated into two parts: a part that can be calculated exactly and which accounts for the majority of the energy, and a small part that acts as a correction term. This correction is needed because KS uses molecular orbitals to represent the electron density. Just as with Hartree–Fock, these orbitals are assumed to be non-interacting. However, in reality there will be interactions, which means there is an error in the model.

The small correction term hereby serves to correct for this error.[e] A general DFT energy expression in KS theory can be written as follows:

$$E_{DFT}[\rho] = T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho]$$

Here, $T_s[\rho]$ is the kinetic energy functional calculated from a single Slater determinant, $E_{ne}[\rho]$ is the nucleus-electron attraction functional, $J[\rho]$ is the Coulomb functional, and $E_{xc}[\rho]$ is the *exchange–correlation functional*. The exchange–correlation (XC) functional contains the remaining part of all electron-electron interactions not covered by the first three functionals, and it is this part of the equation that functions as the correction term.

The problem with the XC functional, however, is that the exact form of this functional is unknown. The other three functionals are known exactly, but the XC functional will have to be approximated. We hereby arrive upon one of the major challenges in DFT, which is to design an XC functional that approximates the unknown exact XC functional. XC functionals are further discussed in the next section.

As mentioned, KS theory is closely related to HF theory, both in its formulation as well as its implementation in various QC programs. Just as with HF, KS-DFT employs an iterative orbital improvement procedure that starts from a trial wave function. Furthermore, KS-DFT is a single-determinant method and uses the same wave function types as Hartree–Fock (restricted, unrestricted, and restricted-open). However, unlike Hartree–Fock, it is able to describe electron correlation and thus does not necessarily require a post-Hartree–Fock-like method to obtain energies close to those corresponding to the exact solution of the Schrödinger equation. It should be noted though that the single-determinant approach may adversely affect results for systems in which static correlation plays an important role. In order to properly treat such systems, several Kohn–Sham-based procedures have been proposed, one example being *spin-restricted ensemble-referenced Kohn–Sham* (REKS).[35] These procedures have currently not yet found widespread use.

### 2.5.2 Exchange–correlation functionals

As mentioned, one of the important goals in the field of DFT is to design an exchange–correlation functional (from here on referred to simply as 'functional',

[e] This error is actually somewhat reminiscent of the missing electron correlation error in Hartree–Fock theory.

as is common in literature) that is as close to the exact, unknown functional as possible. As a result, over the years, many different functionals have been proposed. In earlier years, it was common to design the exchange part and the correlation part of the functional separately, after which these different exchange and correlation functionals could be combined to form various XC functionals. In more recent years, functionals have been proposed for which the exchange and correlation parts were constructed together.

An important aspect of functional design is *parameterization*. Parameters can be included to improve the performance of a functional by optimizing the parameters in such a way that the results with the functional are closer to experimental data. The number of parameters used is different for each functional and depends upon the design philosophy behind its construction. Most functionals use at least a few parameters to improve their performance, but the use of too many may lead to overfitting – a case where the functional works well only for systems related to those included in the benchmark experimental data. A few functionals are also designed to be non-empirical, and can thus be considered to be *ab initio*.[f]

Below follows an overview of some of the widely used functionals, categorized by the fundamental variables they rely on (see also Table 2.5). Such a categorization is referred to as *Jacob's ladder*,[36] where each step higher on the ladder corresponds to an increase in the number of these fundamental variables. The idea behind this metaphor is that each step up the ladder is one step closer to the 'heaven of chemical accuracy' (the exact, unknown functional).[g] A step is hereby often referred to as a *rung*.

[f] It should be noted here that these functionals do contain parameters, however, these have a physical basis (hence, they are non-empirical).

[g] Jacob's Ladder is the ladder to heaven that Jacob dreams about in the Book of Genesis.

***Table 2.5*** *The classification of XC functionals by Jacob's ladder.*

| Rung | Variables | Classification | Examples |
|---|---|---|---|
| 1 | $\rho$ | local density approximation | LSDA, PZ86, VWN |
| 2 | $\rho$, $\nabla\rho$ | GGA | BLYP, PW91, PBE, HCTH, KT2 |
| 3 | $\rho$, $\nabla\rho$, $\nabla^2\rho$ or $\tau$ | meta-GGA | TPSS, $\tau$-HCTH, M06-L |
| 4 | $\rho$, $\nabla\rho$, $\nabla^2\rho$ or $\tau$, exact exchange | hybrid or hybrid (meta)-GGA | B3PW91, B3LYP, PBE0, TPSSh, M05, M06 |
| 5 | $\rho$, $\nabla\rho$, $\nabla^2\rho$ or $\tau$, exact exchange, virtual orbitals | [1] | OEP, B2PLYP, B2GP-PLYP, mPW2PLYP, DSD-PLYP, XYG3 |

[1] There does not seem to be a consensus on a name for the fifth rung. OEP and related methods are often referred to as *generalized random phase approximations*, but the fifth rung is often also associated with double hybrid functionals.

### 2.5.3 The first rung – local density approximation functionals

On the first rung of Jacob's ladder, functionals only use the *local density* as information. It is hereby assumed that this density varies very slowly, making it possible to consider the density at a given point as a uniform electron gas. Examples of widely used exchange functionals on this rung are LDA (local density approximation) and LSDA (local spin density approximation, also referred to as 'Slater' sometimes),[34,37] two very similar functionals that are actually identical for closed-shell systems. Well-known correlation functionals of this rung have been proposed by Vosko, Wilk, and Nusair[38] (VWN, who proposed several functionals), Perdew and Zunger[39] (PZ81), and Perdew and Wang[40] (PW). In general, functionals of this rung are not that accurate for molecular systems, but do give good results for solid state systems (where the electron density is delocalized throughout the solid and a uniform gas assumption is thus appropriate).

### 2.5.4 The second rung – GGA functionals

On the second rung, it is no longer assumed that a uniform gas description is appropriate. This is done by making the functional depend not only on the density, but also the gradient (*i.e.* the first derivative) of the density ($\nabla\rho$). Such functionals are referred to as *generalized gradient approximation* (GGA) functionals. One of the most popular GGA exchange functionals has been proposed by Becke[41] (abbreviated B or B88), and consists of a one-parameter correction (based on the gradient of the density) to LSDA. This single parameter was obtained by fitting against data available for noble gas atoms. Another popular exchange functional is the OPTX functional by Handy and Cohen[42] (O),

which employs two parameters. Correlation functionals of the GGA form have also been proposed. One of the most employed correlation functionals is a four-parameter functional proposed by Lee, Yang, and Parr in 1988[43] (LYP).

A series of non-empirical GGA exchange–correlation functionals has been proposed by Perdew and co-workers. Two of these are PW91[44] (Perdew–Wang 1991), and PBE[45] (Perdew–Burke–Ernzerhof, proposed in 1996). The individual exchange and correlation parts of these functionals have also been used in combination with other exchange or correlation functionals, leading to combinations such as BPW91 and PW91LYP. The exchange part of PW91 has been modified by Adamo and Barone to improve its description of weak interactions, leading to the mPW91 functional.[46]

Other popular GGA functionals are based on the B97 functional, proposed by Becke in 1997. This functional is originally of the fourth-rung (*i.e.* hybrid), but has been reparameterized to a GGA functional by several researchers. Such reparameterizations are B97-D by Grimme[47] (the 'D' in this functional will be explained later on) and the HCTH family of functionals by Handy and co-workers.[48]

A last family of GGA functionals that should be mentioned is the KT family, proposed by Keal and Tozer, which is designed for the calculation of NMR chemical shifts. KT1 and KT2 consist of LDA exchange and VWN correlation plus an additional gradient term,[49] whereby the difference between KT1 and KT2 only lies in their parameters. KT3 consists of LDA and OPTX exchange, LYP correlation, and an additional gradient term.[50] KT3 performs slightly worse than KT1 and KT2 with respect to calculating NMR chemical shielding constants, but is an improvement for other properties, such as atomization energies and reaction barriers.

### 2.5.5 The third rung – meta-GGA functionals

On the third rung are the so-called *meta-GGA functionals* that depend not only on the local density and its first derivative, but also on its second derivative (the *Laplacian*, $\nabla^2\rho$). Alternatively, a meta-GGA functional can depend on the *kinetic energy density* ($\tau$) as this contains the same information. Several meta-GGA functionals have been proposed. $\tau$-HCTH is the $\tau$-dependent member of the HCTH family of functionals.[51] TPSS is a non-empirical meta-GGA functional that can be viewed as the successor to the PBE functional.[52] A final example of a meta-GGA functional is M06-L,[53] which is a pure meta-GGA analogue of the fourth-rung M06 functional.

### 2.5.6 The fourth rung – hybrid functionals

In the case of a system with non-interacting electrons, the unknown, exact XC functional can be reduced to only an exchange functional (as there is no correlation to describe). Furthermore, the *exact* exchange functional for such a system would no longer be unknown – it would actually be identical to Hartree–Fock theory being applied to the KS orbitals. Because of this relation between Hartree–Fock theory and the exchange part of the XC functional, it has been attempted to improve XC functionals by adding a portion of Hartree–Fock theory to them. The resulting functions are called *hybrid* functionals (the term *hyper GGA* is also used). This approach has been proven to be very successful, so much even that the inclusion of this *exact exchange* in XC functionals has become common practice in functional design.

One of the first hybrid functionals is B3PW91, proposed in 1993 by Becke,[54] which has the following form:

$$E_{xc}^{B3PW91} = (1 - a - b)E_x^{LSDA} + aE_x^{B88} + bE_x^{exact} + (1 - c)E_c^{PW\_LDA} + cE_c^{PW91}$$

with a = 0.72, b = 0.20, and c = 0.81 (determined by fitting to experimental data). Here, the exchange part consists of 20 % exact exchange as well as a slight excess of LSDA exchange (recall that B88 consists of LSDA plus a gradient-dependent correction) while the correlation part has a slight excess of PW91 LDA correlation. A well-known variation on this functional, that has surpassed the popularity of the original, is the B3LYP functional in which PW91 is replaced with LYP and PW LDA is replaced with VWN[h]. Despite the emergence of newer, more accurate functionals, B3LYP remains one of the most frequently employed functionals to this day.

Many other hybrid functionals, both hybrid GGA and hybrid meta-GGA, have been proposed. Some of these are hybrid versions of functionals discussed earlier, such as B1PW91, mPW1LYP, PBE0[55] (also known as PBE1PBE) and the recently proposed PBE0-1/3,[56] TPSSh,[52] and τ-HCTH-hybrid.[51] Others are new functionals designed to include a portion of exact exchange. B97, proposed by Becke,[57] is a 10-parameter hybrid GGA functional that has been reparameterized several times (giving, for example, the B98[58] and B97-1[48a] functionals). The M05[59] and M06[60] functionals by Zhao and Truhlar are heavily parameterized hybrid meta-GGA functionals, including 25 and 34 parameters, respectively.

[h] Commonly, the VWN1 RPA or VWN5 functional is used.

### 2.5.7 The fifth rung

Up until this point, only information from occupied KS orbitals has been used. The next rung on Jacob's ladder would be to use information from virtual KS orbitals, similar to that which post-Hartree–Fock methods use. One early attempt at a fifth-rung method are the *optimized effective potential* (OEP) methods,[61] which can be viewed as self-consistent KS-MPx. Experience with OEP (as well as related methods) is as of yet limited, but there are reports that these methods show significant errors even for small systems and are thus probably flawed.[61]

Much more successful are functionals of the *double hybrid* type.[62] Such functionals use an MP2-like term in the correlation part of the functional. One of the first double hybrid functionals proposed is B2PLYP, published by Grimme in 2006,[63] which has the following form:

$$E_{xc}^{B2PLYP} = (1 - a_x)E_x^{B88} + a_x E_x^{exact} + (1 - a_c)E_c^{LYP} + a_c E_c^{PT2}$$

with $a_x$ = 0.53 and $a_c$ = 0.27. It should be noted that this approach is not completely self-consistent. Instead, the KS orbitals are first determined without the MP2-like term (this functional could be denoted as B2LYP), after which the optimized orbitals are subjected to the MP2-like treatment. The results from the perturbation are then added to those obtained with B2LYP, thus yielding the B2PLYP result. After B2PLYP was published, various modifications to the functional were proposed in order to optimize its accuracy. The resulting functionals use different values for $a_x$ and $a_c$ (*e.g.* B2K-PLYP[64] and B2GP-PLYP[65]), different exchange and/or correlation parts (*e.g.* mPW2PLYP[66]), and sometimes use spin-component scaling for the MP2-like term (*e.g.* DSD-BLYP[67] and DSD-PBEP86[68]). Another suggested approach is to use a true MP2-based correction (based on an HF reference) instead of an MP2-like correction based on KS orbitals, examples of such a functional being MC3BB and MC3MPW (the first proposed double hybrid functionals).[69] Finally, it has been suggested to make the double hybrid functional fully non-self-consistent, *i.e.* first determine the KS orbitals with a different functional (such as B3LYP) and then use these orbitals for each term of the double hybrid functional (an example of such a functional being XYG3[70]). In general, the accuracy of double hybrid functionals has been found to be higher than that of hybrid functionals, but this comes at a significantly higher computational cost.

### 2.5.8 Common problems with DFT

If the unknown, exact XC functional would be known, DFT would be an exact method. However, all currently proposed functionals are merely approximations to this exact XC functional. As a result, cases exist where these functionals generally have difficulties in giving an accurate description.

One such a difficult case for current functionals is London-dispersion forces. For example, noble gas atoms in reality show a slight attraction to each other. Many functionals, however, predict that the interaction between such atoms is repulsive. This error is generally caused by the fact that functionals of the first three rungs depend only on the local density and derivatives thereof (and possibly the kinetic energy density). Hybrid functionals and double hybrid functionals usually perform better in such cases, though there is still room for improvement. The dependence on the local density also causes problems in the case of charge-transfer systems. A third example of a difficult area that should be mentioned is systems with loosely bound electrons, such as anions and Rydberg states[i].

Various methods have been proposed to make up for these deficiencies, two of which will be discussed below. These are the *empirical dispersion correction* scheme and the *long-range correction* scheme.

The empirical dispersion correction scheme, proposed by Grimme,[71,47] is a correction that is calculated based on the interatomic distances in the system, hereby using various predetermined coefficients. It does not use information from the KS orbitals, nor adds information to them. A dampening function is employed to suppress the correction at small interatomic distances, as this could lead to singularities or double-counting effects. Various versions of the scheme have been proposed, which are generally referred to as D1, D2, D3, and D3BJ (D3 with a dampening function proposed by Becke and Johnson).

The empirical dispersion correction scheme has been found to be very successful. Many functionals perform better when the correction is used, even for systems in which long-range interactions do not play a significant role.[72] Some functionals were even designed to be used with the correction scheme, whereby parameters of the functional and the correction scheme were optimized simultaneously. Examples of such functionals are the GGA B97-D, SSB-D,[73] and

[i] A Rydberg state is an electronic state in which an electron occupies a highly diffuse orbital.

S12g[74] functionals, the hybrid S12h[74] and ωB97X-D[75] functionals (the latter of which also uses the long-range correction scheme discussed below), and the double hybrid DSD-PLYP functional.

When considering wave function based methods, it has been shown that the importance of dynamic correlation decreases rapidly with increasing electron-electron distances. This implies that at large electron-electron distances Hartree–Fock should be able to provide an accurate description. Based on this rationale, a long-range correction (LC) scheme has been proposed[76] in which the amount of exact exchange in a functional is not constant, but instead depends on the electron-electron distance. Functionals following this scheme are often also referred to as being *range-separated*.

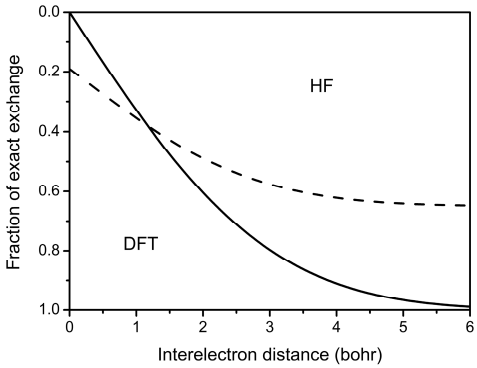


***Figure 2.11*** *The amount of exact exchanged used at various electron-electron distances for the LC-ωPBE (solid) and CAM-B3LYP (dashed) functionals.*

Various LC functionals have been proposed, two of which are considered in Figure 2.11. For the LC-ωPBE functional,[77] which is based on PBE, the amount of exact exchange ranges from 0 to 100 %. This implies that the functional is an almost pure GGA functional at close electron-electron distances while, at long distances, the exchange part of the functional is almost fully exact. For the CAM-B3LYP functional,[78] which is based on B3LYP, the amount of exact exchange ranges from 19 to 65 %, values that were determined by fitting to experimental data. Other examples of functionals that use the LC scheme are ωB97X-D, M11,[79] and a pure meta-GGA version of M11 (M11-L)[80] in which the exact exchange is replaced with a second meta-GGA functional that is optimized for long-range descriptions. It is important to note that it has been found that LC functionals on average do not always perform better than 'normal' hybrid functionals. For calculations on systems in their ground state, they might

actually perform slightly worse. However, for excited state calculations (*i.e.* TDDFT, which will be discussed later on), the performance of LC functionals is usually superior in the case of charge-transfer and Rydberg states.

## 2.6 Time-dependent HF and DFT

So far, a few methods for calculating excited electronic states have been discussed. Configuration interaction (with a single-determinant reference wave function) can be used to calculate the properties of excited states, but does so using ground state orbitals. The related MCSCF-based methods are more advanced, as these methods are able to optimize their orbitals for any state. In addition, these methods can also describe interactions between states by optimizing the orbitals for several states at once.

Another important approach towards calculating excited electronic states is based on calculating the *time-dependent response* of the wave function upon an external perturbation. In its most simple form, such an approach considers merely the *linear response* of a system to the perturbation. This has the benefit that knowledge of the wave function after the perturbation is not needed; only knowledge of the wave function prior to the perturbation (*i.e.* the ground state wave function) is needed.[j] When applied to a Hartree–Fock reference wave function, this approach is referred to as *random phase approximation* (RPA) or *time-dependent Hartree–Fock* (TDHF). When the reference wave function is from a DFT calculation, the approach is called *time-dependent DFT* (TDDFT).[81] When applied to a coupled cluster wave function, the approach is often called *equation-of-motion coupled cluster*.[82]

In general, the accuracy of TDHF is higher than that of the cheaper CIS method (which can be viewed as an approximation to TDHF). TDDFT is often more accurate than TDHF and has become a very popular method for the calculation of vertical excitation energies and oscillator strengths as well as the calculation of excited state geometries. It is however important to note that linear response TDHF and TDDFT do suffer from difficulties in describing the region around conical intersections on the PES.[83] Modifications of the TDDFT method, such as spin-flip TDDFT,[84] have been proposed in order to overcome this limitation.

[j] It should be noted that, in principle, the time-dependent response approach described here is not limited to the calculation of excited states, but can be used for calculating other properties, such as hyperpolarizabilities. Furthermore, the perturbation does not have to be an electric one, it may for example also be magnetic.

Popular functionals for TDDFT often use a large portion of exact exchange, one example being the BH&HLYP functional which consists of 50 % B88 exchange and 50 % exact exchange. This is because such functionals perform better for the description of states with a strong charge-transfer character. It is also for this reason that range-separated functionals such as CAM-B3LYP are a popular choice for TDDFT. Double-hybrid functionals are rarely used for TDDFT though, because the use of such a functional would require a TD-MP2-like calculation which is computationally very expensive. Therefore, it has been proposed to replace the TD-MP2-like calculation with a CIS(D)[k] one.[85] However, this CIS(D) calculation is merely used as a correction to the calculated excitation energies, and does not improve other properties such as transition dipole moments (which are instead calculated using a single hybrid functional).

A last method that should be mentioned is the Tamm–Dancoff approximation to TDDFT, commonly abbreviated as TDDFT/TDA.[86] This method relates to TDDFT as CIS relates to TDHF, and is cheaper than TDDFT. It has also been reported that TDDFT/TDA is more robust than TDDFT for chemically more complex systems.[86,87]

## 2.7 Basis sets

### 2.7.1 Describing molecular orbitals

All of the theories discussed so far have in common that they make use of molecular orbitals. These orbitals are typically described using a combination of known functions, the so-called *basis set*. If an infinite number of these *basis functions* would be used, this description would be fully accurate and the basis set would be *complete*. However, this is not possible for actual calculations and as a result a limited number of functions has to be used. The use of a basis set can therefore be seen as (yet) another approximation that is made in computational chemistry.

We have seen before that the computational cost of various methods increases rapidly with the number of basis functions. As a result, one would like to limit the number of basis functions as much as possible without compromising the accuracy of description of the molecular orbitals too much. For this reason, early computational chemists studying molecular systems made use of *Slater-type*

[k] CIS(D) is a method that applies a perturbative Doubles correction to CIS state energies. Compared to CISD, this method is cheaper as well as more suitable for the calculation of excited states. See: Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee T. J. *Chem. Phys. Lett.* **1994**, *219*, 21–29.

*orbitals* (STOs) that were centered on the nuclei (Figure 2.12). An STO has the advantage that it accurately describes the cusp of the orbital close to the core while providing an exponential decay at longer distances from the core. However, the use of STOs has an important drawback from a mathematical point of view: with STOs, not all calculations can be performed analytically.[1]

The solution to this problem is to replace STOs with *Gaussian-type orbitals* (GTOs). GTOs have a different shape than STOs. However, by combining multiple GTOs, the shape of each STO can largely be reproduced (Figure 2.12). (This practice of combining GTOs into one function is called *contracting*.) One of the first basis sets to use this principle is the STO-3G basis from Pople and co-workers, which uses 3 GTOs to reproduce each STO.[88] Figure 2.12 shows this for the 1*s* orbital of a hydrogen atom. As can be seen, the shape of the STO is approximated well at intermediate to long distances. However, the cusp of the STO at close distances is not reproduced well. Despite this drawback, GTO basis sets have become very popular in the modeling of molecular systems and QC programs that work with STOs are nowadays in the minority.

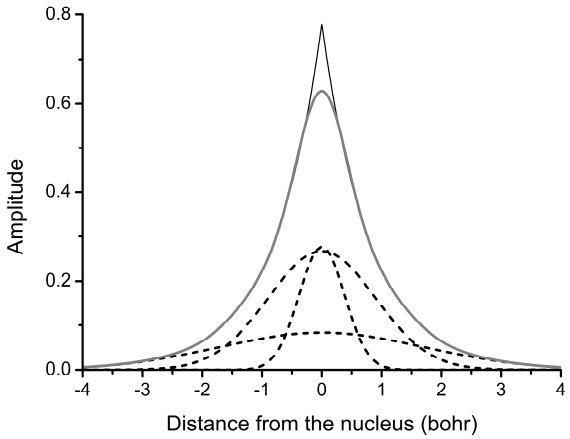


***Figure 2.12*** *The STO-3G basis set for hydrogen. Three GTOs (dashed) are combined to form one function (solid, grey) which approximates an STO (solid, thin).*

### 2.7.2 Anatomy of a basis set

The simplest STO or GTO basis set (with functions centered on the nuclei) uses one contracted function for each atomic orbital. For hydrogen and helium, this

[1] To be specific, for systems consisting of more than two atoms it is not possible to calculate many of the two-electron integrals analytically.

means that a single basis function is employed to create the 1*s* orbitals. For first-row atoms, only five functions are needed: two *s* functions (which provide descriptions for the 1*s* and 2*s* orbitals), and one set of *p* functions (for the $2p_x$, $2p_y$, and $2p_z$ orbitals). For second-row atoms, three *s* functions and two sets of *p* functions are needed. For third-row atoms, four *s* functions, 3 sets of *p* functions, and one set of *d* functions are needed. And so on. Such a basis set is called a *minimal set*.

The quality of the basis set can be improved by doubling all basis functions, which gives a so-called *double zeta* (DZ) basis set. Such a basis set is much more flexible than a minimal set is. However, due to the increase in functions it is also computationally much more expensive. For this reason, the core functions are usually not doubled, which makes sense as chemical processes typically involve only the valence orbitals. The resulting set is often referred to as being *split valence*, though the term DZ is also used for these sets. The quality of the set can be further improved by tripling (TZ) or quadrupling (QZ) the number of basis functions.

A DZ, TZ, or QZ set as described above is still not able to produce accurate molecular orbitals for the majority of systems. Specifically, it lacks the flexibility needed to allow orbitals to exhibit a certain degree of asymmetry. The flexibility can be increased by adding *polarization functions*, which are higher angular momentum functions used for polarizing functions of one momentum lower. For example, an *s* function can be polarized by a *p* function, a *p* function can be polarized by a *d* function, a *d* function can be polarized by a *f* function, and so on (see Figure 2.13). Multiple polarization functions of various angular momentums can be added to a basis set this way.

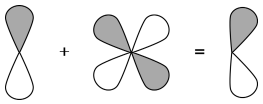


***Figure 2.13*** *A* p *function that is polarized by a* d *function.*

This raises the following question: what is the optimal amount for polarization functions for a basis set? There is no straightforward answer to this, but for main group elements many modern basis sets typically consist of one less set of functions with each increase in angular momentum. For example, a TZ set is usually of the 3*s*2*p*1*d* type for hydrogen, and of the 4*s*3*p*2*d*1*f* type for carbon. Sets for transition metals usually follow this form approximately (as they contain more *d* functions in order to describe the *d* valence orbitals). Sets that

have far less polarization functions, such as 4*s*3*p*1*d*, can be considered to be underpolarized, which may have an adverse effect on the accuracy. Conversely, a set such as 3*s*2*p*2*d*2*f* can be considered to be overpolarized.

Many basis sets, both GTO and STO, do not provide a good description of the electron density far from the nuclei. This originates from a lack of *diffuse* (*i.e.* broad) functions being present in the set. Typically, the absence of such functions hardly affects the accuracy. However, in some cases, an accurate description of the electron density at long distances *is* important, for example when studying van der Waals interactions. For such cases, a basis set is typically *augmented* with diffuse functions which are kept uncontracted (*i.e.* they are kept as separate functions) as this increases the flexibility of the set. When a set has only diffuse functions of the *s* and *p* angular momentums it is said to be *minimally augmented*. Some sets on the other hand employ diffuse functions of all angular momenta that are already in the set (causing 4*s*3*p*2*d*1*f* to become 5*s*4*p*3*d*2*f*).

### 2.7.3 Effective core potentials

When the molecular system being studied contains heavy elements, the use of a basis set as described above has two important drawbacks. First of all, heavy elements consist of many orbitals which means that a large number of basis functions is needed to describe the element properly. However, typically the core orbitals of an element are chemically not interesting, so this is not very economical in terms of computational resources. Secondly, for heavy elements relativistic effects play an important role, and these cannot easily be described.[m]

A solution to these drawbacks is to replace the basis functions that describe the core orbitals of an element with a function that is designed to model the element's core. Various such functions have been suggested, but the most popular of them is the *effective core potential* (ECP, also referred to as *pseudopotential*, PP). ECPs as implemented in most QC programs can account for the scalar relativistic effects while some of the more recent implementations are also able to include spin-orbit effects. In the past decades, various ECPs have been suggested, some even for light elements (solely for the purpose of reducing computational cost).

### 2.7.4 Popular basis sets

The text below discusses a few popular families of basis sets.

[m] Methods capable of calculating relativistic effects are beyond the scope of this text.

*The STO-3G, 3-21G, 6-31G, and 6-311G sets*

The STO-3G, 3-21G, 6-31G, and 6-311G sets by Pople and co-workers are very popular as they use a relatively small amount of basis functions and are thus quite economical. The STO-3G set, where 3G refers to the amount of Gaussian functions in the contraction, was designed to approximate the shape of STOs as well as possible. The 3-21G set is a cheap double zeta set that uses 3 Gaussian functions for the core and 2 + 1 Gaussian functions for the valence orbitals.[89] Instead of trying to approximate STOs, this set was designed to give as low energies as possible in HF calculations (using the variational principle to optimize the basis set). The 6-31G set[90] and the 6-311G set[91] (which is a triple zeta set) follow the same design principle. It should be noted that these sets all use combined *sp* functions for the description of the valence orbitals instead of separate *s* and *p* functions. Also, the 6-311G set for second-row atoms was proposed by McLean and Chandler.[91b]

Polarization functions are available for the 6-31G and 6-311G sets and are indicated in brackets, *e.g.* 6-311G(2df,2p), where the first part ('2df') denotes polarization functions on Li and heavier, and the second part ('2d') denotes polarization functions on H and He. For the 3-21G set, a single set of *d* functions is available only for second-row atoms. Diffuse functions are available for the 3-21G, 6-31G, and 6-311G sets and are denoted by '+' signs (such as 6-31++G). '+' indicates diffuse functions are added to Li and heavier. '++' indicates diffuse functions are added also to H and He. These diffuse functions are of the *s* and *sp* type only (*i.e.* minimal augmentations).

*Dunning's correlation consistent cc-pVXZ sets*

When the size of a basis set grows from minimal to DZ to TZ to QZ and so on, its results will come increasingly close to that obtained with a (hypothetical) complete basis set (CBS). However, the speed with which a system will converge towards the CBS limit depends on the method used. Specifically, HF and KS-DFT converge much faster than post-HF methods. The *correlation consistent* sets by Dunning and co-workers were designed to provide rapid as well as smooth convergence towards the CBS limit for post-HF methods by recovering as much electron correlation as possible.[92] In order to achieve this, the *s* and *p* functions were optimized using HF while the polarization functions were optimized at the CISD level. The resulting sets are known as cc-pVDZ, cc-pVTZ, cc-pVQZ, cc-pV5Z, etc. Polarization functions are included in each set while diffuse functions can be added separately (and are denoted by an 'aug-' prefix).

Some important variations on the cc-pVXZ sets should be mentioned. For the elements Al – Ar, the cc-pV(X+d)Z sets are available which contain an extra set of *d* functions designed to improve the smoothness of the convergence towards the CBS limit (important if one would like to extrapolate the results towards CBS).[93] The cc-pVXZ-PP sets, available for most elements in the range Cu – Rn, contain ECPs that describe the core orbitals and include relativistic effects.[94] Finally, the cc-pwCVXZ sets should be mentioned,[95] which contain extra functions in order to improve the description of core-valence electron interactions (typically, basis sets are designed primarily for describing valence-valence interactions).

*Jensen's polarization consistent pc-X sets*

The correlation consistent basis sets provide rapid and smooth convergence for post-HF methods. However, for HF and KS-DFT calculations their convergence towards the CBS may be suboptimal. Jensen has proposed a family of sets that is based on the correlation consistent sets and that is optimized using HF calculations only: the *polarization consistent* sets.[96] The smallest is pc-0, which is a small DZ basis set without polarization functions. The pc-1, pc-2, pc-3, and pc-4 sets are of the DZ, TZ, QZ, and 5Z type, respectively, all with polarization functions included. As with the correlation consistent sets, diffuse functions can be added separately and are denoted with the 'aug-' prefix. Two variations on the pc-X sets are the pcS-X sets[97] and the pcJ-X sets,[98] which are designed for the calculation of NMR shielding constants and spin-spin coupling constants, respectively.

*Ahlrichs' and Weigend's SVP and XZV sets*

Ahlrichs, Weigend, and co-workers have designed various sets of the DZ, TZ, and QZ type. The first of these are def-SV(P), def-TZVP, and def-QZVP.[99] Polarization functions are included in the design, but their number can be increased to achieve results closer to the CBS limit, yielding the def-SVP, def-TZVPP, and def-QZVPP sets. In 2005, these sets where updated to improve their convergence towards the CBS limit as well as include ECPs for the elements Rb – Rn.[100] These sets are denoted by the prefix 'def2-'. In turn, the 'def2-' sets for Rb – Rn were recently improved to give the 'dhf-' sets.[101] None of these sets contain diffuse functions, nor have Ahlrichs, Weigend, and co-workers proposed any such functions. However, partial augmentations optimized for the calculation of molecular response properties have been proposed by Rappoport and Furche[102] (the SV(P)D/SVPD/XZVPD/XZVPPD sets) while general purpose minimal augmentations have been proposed by Truhlar and co-workers (the 'ma-' sets).[103] It is also common practice to augment the def2/dhf sets with diffuse functions from the correlation consistent sets described above.

## 2.8 Studying molecular systems

So far, this chapter has mainly focused on explaining the theory behind several methods and how methods relate to each other. In this section, various common practices in computational studies on organic molecules are described.

### 2.8.1 Geometry optimizations towards a minimum

Comparable to the SCF procedure in Hartree–Fock theory, the location of a *minimum* (a 'stable' geometry) on the potential energy surface is a step-wise, iterative process. For the optimization of molecular geometries, a quasi (or pseudo) Newton–Raphson method is most commonly used. In order to explain quasi Newton–Raphson optimizations, it is necessary to introduce two concepts. First of all, the optimization in each step requires the calculation of the *gradient*, which is the first derivative of the energy on a certain point on the PES. Also needed is the second derivative of the energy, which is called the *Hessian matrix* (or *force constant matrix*). The gradient and Hessian are both used by the optimization algorithm to 'drive' the starting geometry across the PES towards its intended goal, *e.g.* the nearest minimum. Figure 2.14 shows a schematic representation of a quasi Newton–Raphson optimization.

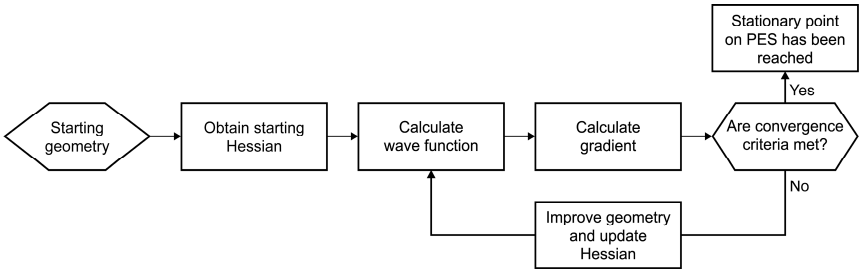


***Figure 2.14*** *Simplified representation of a quasi Newton–Raphson geometry optimization procedure.*

Starting from a certain geometry, the first step in the optimization procedure is to obtain the Hessian. The most accurate would be to calculate the Hessian at the level at which the optimization is performed, *e.g.* at the RHF/6-31G(d) level. However, this is typically quite expensive. Fortunately, the optimization procedure doesn't require the precise Hessian. Instead, it may also be approximated, though this usually causes the whole procedure to require more iterations to converge. The cheapest is hereby to 'guess' a Hessian, and this is done most frequently. When a Hessian has been obtained, the next steps are the

calculation of the energy of the system (for RHF this is done using the SCF procedure discussed earlier) followed by the calculation of the gradient.[n]

When the gradient is known, it is checked to see if a stationary point on the PES has been reached. If this is not the case, the optimization algorithm will use the calculated gradient and Hessian to determine a new geometry. Next, with a new geometry known it is possible to recalculate the Hessian. However, it is often a better choice to update the Hessian, because a) if the starting Hessian was guessed, the updated Hessian is actually closer to exact Hessian than a new guess, and b) if the exact Hessian was calculated, it is almost always too expensive to recalculate it for every step. Finally, with a new geometry and Hessian known, a next iteration can be performed, starting with calculating the energy.

### 2.8.2 Vibrational analysis and thermochemistry

When the optimization procedure has fulfilled the convergence criteria, a stationary point on the PES has been found. However, such a point does not necessarily correspond to a minimum. This is because the convergence criteria typically only require the gradient to become close to zero. This is the case at the bottom of a minimum, but also when on a saddle point.

To investigate the nature of the stationary point found, one can calculate the Hessian at the same level used for the geometry optimization. In other words, if the geometry optimization was for example performed at the B3LYP/6-31G(d,p) level, the calculation of the Hessian should be done at the same level. The calculated Hessian can then be used to produce a list of vibrational modes that provides information about the stationary point found. If this point should correspond to a minimum, the list should not contain any imaginary vibrational modes (often printed as modes with a negative wavenumber). Furthermore, the wavenumbers of the six modes[o] that correspond to rotations and translations of the system should be close to zero.

When the above analysis shows that the stationary point is indeed a minimum, the vibrational data provided by the Hessian may be used for calculating thermochemical data. The energies calculated by all theories discussed so far have in common that they provide the energy of the system in absence of any

[n] If the exact Hessian was calculated, these would already be known for the current geometry.

[o] Five modes if the molecule is linear.

atomic movement. This is, however, not realistic – atoms are never motionless, not even at 0 K. As a result, the real energy of the system lies above the calculated energy. The energy of the system at 0 K, the so-called *zero-point energy*, is as follows:

$$E_{ZPE} = E_{SCF} + E_{vib,0K}$$

where $E_{SCF}$ is the calculated energy (called the *SCF energy* for SCF methods) and $E_{vib,0K}$ is the vibrational contribution to the zero-point energy. At T > 0 K, the thermal energy of the system is as follows:

$$E_{thermal} = E_{ZPE} + E_{rot} + E_{trans} + E_{vib}$$

where $E_{rot}$, $E_{trans}$, and $E_{vib}$ are the rotational, translational, and vibrational contributions to the energy at a given temperature. The corresponding enthalpy is as follows:

$$H = E_{thermal} + RT$$

In addition, the Hessian can be used to obtain the entropy of the system, and thus allows one to calculate the system's Gibbs free energy.

It should be noted that vibrational data from the Hessian uses an important approximation, namely that all vibrations are harmonic. However, in reality vibrations are anharmonic, as is shown in Figure 2.15 for the stretching vibration of a dihydrogen molecule. Because of this, vibrational data from the Hessian has a (typically minor) error that is also present in the thermochemical data. The thermochemistry can be improved by scaling the vibrational contributions to the energy by a predetermined factor or by the use of an *ad hoc* correction scheme (several have been proposed).[104] In addition, it is possible to obtain vibrational data from higher-order derivatives.

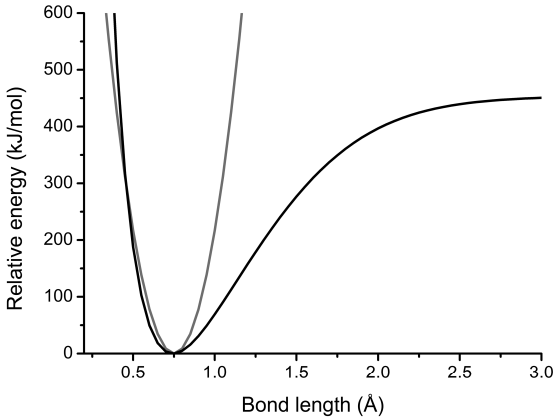


***Figure 2.15*** *Potential energy curve of the stretching vibration in the dihydrogen molecule, calculated at the full CI/aug-cc-pV5Z level. Shown are the harmonic potential curve (gray) and the true, anharmonic potential curve (black).*

### 2.8.3 Optimization towards a saddle point

The procedure for a geometry optimization towards a first-order saddle point (*i.e.* a transition state) does not differ much from that of an optimization towards a minimum. The flowchart of Figure 2.14 still applies, but there is an important difference with respect to the Hessian. As was discussed, for a true minimum all vibrational modes will have a frequency with a positive sign. However, on a first-order saddle point, there will be one vibrational mode with a negative sign, a so-called *imaginary* mode. The nature of this mode corresponds to the process to which the transition state belongs. For example, the transition state of a bond breaking process will exhibit an imaginary mode in which the bond to be broken will be stretching.

A question that might arise here is how the optimization algorithm knows which saddle point one is looking for. For quasi Newton–Raphson methods, the answer is that it does not know, and as a result, one will have to provide a starting geometry already close to the saddle point.[p] Most importantly, in the Hessian obtained for this geometry the desired imaginary mode should already be present. A guessed Hessian cannot contain any imaginary modes, so one will have to either calculate the precise Hessian or use another (possibly cheaper) QC method to obtain a Hessian with the required imaginary mode. Starting from a

[p] This is also true for optimizations towards a minimum, but it is much more important for saddle point optimizations as these are typically difficult to locate.

reasonable geometry and Hessian, the optimization algorithm will then try to maximize the desired imaginary mode while making any other modes positives.

Once a saddle point is found, its nature can be investigated by calculating the precise Hessian. Only one imaginary mode may be present and, as before, the rotational and translational modes should be (close to) zero. In addition, one can investigate whether the transition state found connects the correct two minima (*i.e.* the reactant and product). This can be done with an *intrinsic reaction coordinate* (IRC) calculation, in which the atomic displacements in the imaginary mode are 'followed' to the closest minima.

### 2.8.4 Vibrational spectroscopy

Because the vibrational modes obtained by calculating the Hessian correspond to experimentally observed molecular vibrations, they can be used to predict or explain experimentally obtained vibrational spectra. However, while the Hessian contains information about the frequency of each mode, it does not contain any information with respect to the activity of each mode. Thus, these have to be calculated separately.

Infrared spectroscopic data is typically cheap to obtain. A vibrational mode is IR active only when the vibration causes a change in the molecule's electric dipole moment. As such, obtaining IR activities requires calculating the derivative of the electric dipole moment with respect to each vibrational mode. Because this has a low cost once the Hessian has already been calculated, many QC programs calculate IR activities right after obtaining the Hessian. Often, each IR activity is convoluted with a Lorentzian or Gaussian curve in order to obtain a spectrum with broadened peaks that resembles the experimental spectrum better (Figure 2.16).

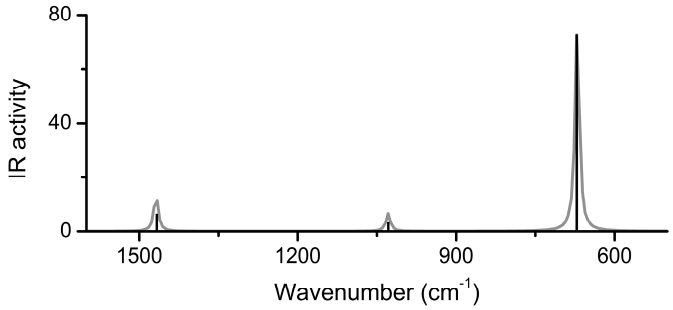


***Figure 2.16*** *IR activities (black) of benzene calculated at the B3LYP/6-31G(d,p) level. The spectrum was plotted using Lorentzian curves (gray).*

Calculating Raman activities is typically more expensive than calculating IR activities. A vibrational mode is Raman active when the vibration causes a change in the molecule's polarizability. This requires calculating the derivative of the polarizability with respect to an external electric field, which is more complex than obtaining the derivative of the dipole moment (needed for calculating IR activities). The Raman activities obtained this way are independent of the excitation wavelength and might differ from experimentally observed Raman intensities. However, one can convert the calculated activities to Raman intensities using the formula:[105]

$$I_i = \frac{f(\nu_0 - \nu_i)^4 S_i}{\nu_i(1 - e^{-\frac{hc\nu_i}{kT}})}$$

where $I_i$ is the Raman intensity, $S_i$ is the Raman activity ($Å^4$ $amu^{-1}$), $\nu_i$ is the wavenumber ($cm^{-1}$) of the *i*th vibration, $\nu_0$ is the wavenumber ($cm^{-1}$) of the excitation laser, h is the Planck constant (J s), c is the speed of light (m $s^{-1}$), k is the Boltzmann constant (J $K^{-1}$), T is the temperature (K), and f is an optional normalization factor that is applied to all peaks. Methods also exist for calculating resonance Raman spectra, though these are typically much more complex as they require knowledge of excited electronic states of the system.

Because of the harmonic approximation as well as inaccuracies in the theory and basis set used, calculated wavenumbers will almost always be at too high compared to the wavenumbers observed experimentally. For this reason, it is common practice to scale the wavenumbers by a correction factor if one would like to compare the theoretical and experimental results. Correction factors have been published for various combinations of QC theories (including various XC functionals) and basis sets. One list of factors can be found in the Computational Chemistry Comparison and Benchmark Database.[106]

Finally, it should be noted that calculated Raman activities can depend strongly on the quality of the basis set. Especially the presence of diffuse functions in the set has been found to be important,[107] while the size of the basis set (*i.e.* double zeta, triple zeta, etc.) is of lesser importance.[108] Wavenumbers and IR activities, on the other hand, seem to benefit little from the addition of diffuse functions and depend more on the size of the basis set.[108]

### 2.8.5 NMR spectroscopy

Various methods exist for the calculation of NMR shielding constants. One of the most popular ones is the *gauge including/independent atomic orbitals* (GIAO) method, which is implemented in various QC programs. Some programs also allow for the calculation of spin-spin coupling constants, though this requires a significant additional computational cost. As was discussed earlier, basis sets exists which are designed for the calculation of NMR properties, such as Jensen's pcS-X and pcJ-X sets. Recommending a QC method is however less straightforward as some studies into the quality of calculated NMR data contradict each other in their recommendations.[109,110] However, good choices seem to be the KT1,[109,111] KT2,[109,111] WP04/WC04,[q,110] and OPBE[112] XC functionals as well as the post-Hartree–Fock MP2 method.[109] In addition, it is recommended to employ a solvation model, especially if one is interested in chemical shifts for hydrogen atoms.[110] It should be noted that current implementations typically have problems describing shielding constants in systems containing heavy elements (such as bromobenzene, for which the shielding constant of the carbon attached to the bromine will be predicted to be too low).

## 2.9 Calculation details

**Figure 2.5 – Dissociation curves for H2**

RHF and UHF calculations were performed with Gaussian 09 C.01.[113] The 'exact solution' curve was obtained from a Full CI calculation, performed with Firefly 8.0.0 beta[114] QC program, which is based partially on GAMESS US source code.[115] All calculations used an aug-cc-pV5Z basis set.

**Figure 2.8 – Energies of LiF calculated with MP methods**

The LiF geometry was optimized at the MP2/cc-pVTZ level. MPX energies were calculated with a cc-pVTZ basis set. The calculations were performed with Gaussian 09 C.01.

**Figure 2.10 – Dissociation curves for LiF**

SA-CASSCF energies were obtained with a (6,6) active space consisting of the Li 2*s* and F 2$p_x$, 2$p_y$ 2$p_z$, 2$p_{x'}$, and 2$p_{y'}$ orbitals and with state averaging over the lowest two roots. The SA-CASSCF orbitals were subsequently used as the reference for the XMCQDPT2 and MRCISD calculations. XMCQDPT2 calculations used a 2 × 2 sized model space and an energy denominator shift of

[q] The WP04 and WC04 functionals are reparameterized versions of the B3LYP functional and are optimized to predict hydrogen and carbon chemical shifts, respectively, in chloroform.

0.02 Hartree. SA-CASSCF and XMCQDPT2 calculations were performed with Firefly 8.0.0 beta. MRCISD calculations were performed with the ORMAS program in GAMESS US. All calculations used a DZV basis proposed by Bauschlicher and Langhoff.[116]

**Figure 2.15 – Potential energy curves for H2**
The 'anharmonic' curve is identical to the 'exact solution' curve of Figure 2.5. The 'harmonic' curve was obtained from a Hessian calculation at the same level (bond length: 0.74166 Å; force constant: 1151.63 N $m^{-1}$).

**Figure 2.16 – Potential energy curves for H2**
Calculation was performed with Gaussian 09 C.01. The convoluted spectrum was plotted with GaussSum 2.2.5 using Lorentzian curves with the full width at half maximum (FWMH) set to 10 $cm^{-1}$.

## 2.10 References

[1] Born, M.; Oppenheimer, J. R. *Ann. Phys. (Leipzig)* **1927**, *84*, 457–484.
[2] Born, M.; Fock, V. A. *Z. Phys. A* **1928**, *51*, 165–180.
[3] Slater, J.; Verma, H. C. *Phys. Rev.* **1929**, *34*, 1293–1295.
[4] Berthier, G. *C. R. Acad. Sci. Paris* **1954**, *238*, 91–93.
[5] Pople, J. A.; Nesbet, R. K. *J. Chem. Phys.* **1954**, *22*, 571–572.
[6] Roothaan, C. C. J. *Rev. Mod. Phys.* **1960**, *32*, 179–185.
[7] Sherrill, C. D.; Schaefer, H. F. *Adv. Quant. Chem.* **1999**, *34*, 143–269.
[8] Møller, C. ; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.
[9] Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.
[10] Grimme, S.; Goerigk, L.; Fink, R. F. *WIREs Comput. Mol. Sci.* **2012**, *2*, 886–906.
[11] Hill J. G., Platts, J. A. *J. Chem. Theory Comput.* **2007**, *3*, 80–85.
[12] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793–9802.
[13] Grimme, S. *J. Comput. Chem.* **2003**, *24*, 1529–1537.
[14] Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697–1708.
[15] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
[16] Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. *Chem. Phys. Lett.* **1994**, *219*, 21–29.
[17] Brueckner, K. A. *Phys. Rev.* **1954**, *96*, 508–516.
[18] Handy, N. C.; Pople, J. A.; Head-Gordon, M.; Raghavachari, K.; Trucks, G. W. *Chem. Phys. Lett.* **1989**, *164*, 185–192.
[19] Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968–5975.
[20] Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157–173.
[21] a) Ruedenberg, K.; Sundberg, K. R. in *Quantum Science*; Calais, J.-L.; Goscinski, O.; Linderberg, J.; Öhrn, Y., Eds.; Plenum, New York, 1976, 505–511. b) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41–49. c) Ruedenberg,

K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 51–64. d) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 65–78.
[22] Roos, B. in *Ab Initio Methods in Quantum Chemistry, II*; Lawley, K. P., Ed; Wiley, New York, 1987, 399–446.
[23] Malmqvist, P.-Å.; Rendell, A.; Roos. B. O. *J. Phys. Chem.* **1990**, *94*, 5477–5482.
[24] Nakano, H.; Hirao, K. *Chem. Phys. Lett.* **2000**, *317*, 90–96.
[25] Ivanic, J. *J. Chem. Phys.* **2003**, *119*, 9364–9376.
[26] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
[27] Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374–380.
[28] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger T.; Malrieu, J. P. *J. Chem. Phys.* **2001**, *114*, 10252–10264.
[29] Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299–306.
[30] Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983–7992.
[31] Angeli, C.; Borini, S.; Cestari, M.; Cimiraglia, R. *J. Chem. Phys.* **2004**, *121*, 4043–4049.
[32] Granovsky, A. A. *J. Chem. Phys.* **2011**, *134*, 214113.
[33] Shiozaki, T.; Györffy, T.; Celani, P. Werner, H.-J. *J. Chem. Phys.* **2011**, *135*, 081106.
[34] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.
[35] a) Filatov, M.; Shaik, S. Chem. *Phys. Lett.* **1999**, *304*, 429–437. b) Filatov, M. *WIREs Comput. Mol. Sci.* **2015**, *5*, 146–167.
[36] Kurth, S.; Perdew, J. P.; Blaha, P. *Int. J. Quant. Chem.* **1999**, *75*, 889–909.
[37] Hohenberg P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.
[38] Vosko, S. J.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
[39] Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048–5079.
[40] Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244–13249.
[41] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
[42] Handy N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403–412.
[43] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
[44] Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D.J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671–6687.
[45] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
[46] Adamo C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664–675.
[47] Grimme, S. *J. Comp. Chem.* **2006**, *27*, 1787–1799.
[48] a) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Chem. Phys.* **1998**, *109*, 6264–6271. b) Boese, A. D.; Doltsinis, N. L.; Handy, N. C.; Sprik. M. *J. Chem. Phys.* **2000**, *112*, 1670–1678. c) Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2001**, *114*, 5497–5503. d) Boese, A. D.; Chandra, A.; Martin, J. M. L.; Marx, D. *J. Chem. Phys.* **2003**, *119*, 5965–5980.
[49] Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2003**, *119*, 3015–3024.
[50] Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2004**, *121*, 5654–5660.
[51] Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2002**, *116*, 9559–9569.
[52] Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
[53] Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.
[54] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
[55] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6169.

[56] Guido, C. A.; Brémond, E.; Adamo, C.; Cortona, P. *J. Chem. Phys.* **2013**, *138*, 021104.
[57] Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554–8560.
[58] Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624–9631.
[59] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123*, 161103.
[60] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
[61] a) Talman, J. D.; Shadwick, W. F. *Phys. Rev. A* **1976**, *14*, 36–40. b) Bonetti, A. F.; Engel, E.; Schmid, R. N.; Dreizler, R. M. *Phys. Rev. Lett.* **2001**, *86*, 2241–2244. c) Hesselmann, A.; Gotz, A. W.; Della Sala, F.; Görling, A. *J. Chem. Phys.* **2007**, *127*, 054102.
[62] Zhang, I. Y.; Xu, X. *A New-Generation Density Functional: Towards Chemical Accuracy for Chemistry of Main Group Elements*; SpringerBriefs in Molecular Science; Springer: 2014.
[63] Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.
[64] Tarnopolsky, A.; Karton, A.; Sertchook, R.; Vuzman, D.; Martin, J. M. L. *J. Phys. Chem. A* **2008**, *112*, 3–8.
[65] Karton, A.; Tarnopolsky, A.; Lamère, J.-F.; Schatz, G. C.; Martin, J. M. L. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.
[66] Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.
[67] Kozuch, S.; Gruzman, D.; Martin, J. M. L. *J. Phys. Chem. C* **2010**, *114*, 20801–20808.
[68] Kozuch, S.; Martin, J. M. L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.
[69] Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 4786–4791.
[70] Zhang, I. Y.; Xu, X.; Jung, Y.; Goddard III, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 19896–19900.
[71] a) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473. b) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104. c) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
[72] a) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291–309. b) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465. c) Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.
[73] Swart, M.; Solà, M.; Bickelhaupt, F. M. *J. Chem. Phys.* **2009**, *131*, 094103.
[74] Swart, M. *Chem. Phys. Lett.* **2013**, *580*, 166–171.
[75] Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
[76] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.
[77] Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
[78] Yanai, T.; Tew, D. P., Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.
[79] Peverati, R.; Truhlar, D. G. *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817
[80] Peverati, R.; Truhlar, D. G. *J. Phys. Chem. Lett.* **2012**, *3*, 117–124.
[81] Furche, F.; Burke, K. *Annu. Rep. Comput. Chem.* **2005**, *1*, 19–30.
[82] Koch, H.; Jørgensen, P. *J. Chem. Phys.* **1990**, *93*, 3333–3344.
[83] Harabuchi, Y.; Maeda, S.; Taketsugu, T.; Minezawa, N.; Morokuma, K. *J. Chem. Theory Comput.* **2013**, *9*, 4116–4123.
[84] Shao, Y.; Head-Gordon, M.; Krylov, A. I. *J. Chem. Phys.* **2003**, *118*, 4807–4818.
[85] Grimme, S.; Neese, F. *J. Chem. Phys.* **2007**, *127*, 154116.
[86] Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291–299.
[87] Wang, Y.-L.; Wu, G.-S. *Int. J. Quantum Chem.* **2008**, *108*, 430–439.
[88] a) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657–2664. b) Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 2769–2773.

[89] a) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939–947. b) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1983**, *104*, 2797–2803.
[90] a) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261. b) Dill J. D.; Pople, J. A. *J. Chem. Phys.* **1975**, *62*, 2921–2923. c) Francl, M. M.; Petro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
[91] a) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654. b) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639–5648.
[92] a) Dunning, Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023. b) Woon, D. E.; Dunning, Jr., T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
[93] Dunning, Jr., T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244–9253.
[94] a) Peterson, K. A. *J. Chem. Phys.* **2003**, *119*, 11099–11112. b) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. *J. Chem. Phys.* **2003**, *119*, 11113–11123.
[95] Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **2002**, *117*, 10548–10560.
[96] a) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113–9125. b) Jensen, F. *J. Chem. Phys.* **2002**, *116*, 7372–7379. c) Jensen, F.; Helgaker, T. *J. Chem. Phys.* **2004**, *121*, 3463–3470. d) Jensen, F. *J. Phys. Chem. A* **2007**, *111*, 11198–11204.
[97] Jensen, F. *J. Chem. Theory Comput.* **2008**, *4*, 719–727.
[98] Jensen, F. *Theor. Chem. Acc.* **2010**, *126*, 371–382.
[99] a) Schaefer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577. b) Schaefer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.
[100] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
[101] Weigend, F.; Baldes, A. *J. Chem. Phys.* **2010**, *133*, 174102.
[102] Rappoport, D.; Furche, F. *J. Chem. Phys.* **2010**, *133*, 134105.
[103] Zheng, J.; Xu, X.; Truhlar, D. G. *Theor. Chem. Acc.* **2011**, *128*, 295–305.
[104] a) Pitzer, K. S.; Gwinn, W. D. *J. Chem. Phys.* **1942**, *10*, 428–440. b) Truhlar, D. G. *J. Comput. Chem.* **1991**, *12*, 266–270. c) Ayala P. Y., Schlegel, H. B. *J. Chem. Phys.* **1998**, *108*, 2314–2325.
[105] a) Polavarapu, P. L. *J. Phys. Chem.* **1990**, *94*, 8106–8112. b) Keresztury, G.; Holly, S.; Varga, J.; Besenyei, G.; Wang, A. Y.; Durig, J. R. *Spectrochim. Acta* **1993**, *49A*, 2007–2026.
[106] http://cccbdb.nist.gov/vsf.asp
[107] a) Halls, M. D.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111*, 8819–8824. b) Rappoport, D.; Furche, F. *J. Chem. Phys.* **2007**, *126*, 201104.
[108] Cheeseman, J. R.; Frisch, M. J. *J. Chem. Theory Comput.* **2011**, *7*, 3323–3334.
[109] Kupka, T.; Stachów, M.; Nieradka, M.; Kaminsky, J.; Pluta, T. *J. Chem. Theory Comput.* **2010**, *6*, 1580–1589.
[110] a) Wiitala, K. W.; Hoye, T. R., Cramer, C. J. *J. Chem. Theory Comput.* **2006**, *2*, 1085–1092. b) Jain, R.; Bally, T.; Rablen, P. R. *J. Org. Chem.* **2009**, *74*, 4017–4023.
[111] Keal, T. W.; Tozer, D. J.; Helgaker, T. *Chem. Phys. Lett.* **2004**, *391*, 374–379.
[112] Zhang, Y.; Wu, A. Xu, X.; Yan, Y. *Chem. Phys. Lett.* **2006**, *421*, 383–388.
[113] Gaussian 09, Revision C.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.;

Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A. ; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2010.

[114] Granovsky, A. A.; Firefly version 8.0.0 beta, http://classic.chem.msu.su/gran/firefly/index.html

[115] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

[116] Bauschlicher, C. W.; Langhoff, S. R. *J. Chem. Phys.* **1988**, *89*, 4246–4254.